Contribution à l'élaboration d'un modèle de raisonnement à partir de cas pour l'aide à l'interprétation d'organisations spatiales agricoles

Jean-Luc Metzger

▶ **To cite this version:**

Jean-Luc Metzger. Contribution à l'élaboration d'un modèle de raisonnement à partir de cas pour l'aide à l'interprétation d'organisations spatiales agricoles. Autre [cs.OH]. Université Henri Poincaré - Nancy 1, 2005. Français. NNT : 2005NAN10018 . tel-01754386

HAL Id: tel-01754386
https://hal.univ-lorraine.fr/tel-01754386

Submitted on 30 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Département de formation doctorale en informatique
UFR STMIA

École doctorale IAEM Lorraine

Contribution à l'élaboration d'un modèle de Raisonnement à Partir de Cas pour l'aide à l'interprétation d'organisations spatiales agricoles

THÈSE

présentée et soutenue publiquement le 01-04-2005

pour l'obtention du

Doctorat de l'Université Henri Poincaré – Nancy 1

(spécialité informatique)

par

Jean-Luc Metzger

Composition du jury

Président :	Jean-Paul Haton
Rapporteurs :	Jérôme Euzenat Alain Mille
Directeurs :	Sylvie Lardon Amedeo Napoli
Examinatrice :	Florence Le Ber
Invité :	Jean-Marc Meynard

S.C.D. - U.H.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Mis en page avec la classe thloria.

Remerciements

Je tiens à remercier les personnes qui m'ont fait l'honneur de participer à mon jury et de s'intéresser à ce travail. Je remercie tout particulièrement Jean-Paul Haton, Professeur à l'Université Henri Poincaré, qui a accepté de présider le jury et Jean-Marc Meynard, chef du département SAD de l'INRA, d'avoir accepté l'invitation à participer à ce jury.

Le comité de pilotage, institué par l'INRA, aura été pour moi l'occasion de présenter l'avancement de mon travail. Je remercie ses membres : Michel Duru, Jérôme Euzenat, Bernard Hubert, Alain Mille, Pascal Monestier, Claudine Thenail ainsi que mes encadrants pour le temps qu'ils m'ont consacré et pour leurs conseils et remarques qui ont été riches et qui m'ont permis de faire le point une fois par an pour m'orienter dans la bonne direction. Jérôme et Alain ont de plus été les rapporteurs de cette thèse.

Je tiens à remercier Sylvie Lardon et Amedeo Napoli d'avoir accepté d'encadrer ma thèse à la suite de mon stage de DEA. Je tiens à remercier tout particulièrement Florence Le Ber, pour sa patience et son suivi de mon travail. Sans ses relectures, conseils et soutien ce travail n'aurait probablement pas vu son aboutissement. Merci à tout les trois pour avoir été les examinateurs de mon jury.

Je tiens aussi à remercier les chercheurs du SAD, notamment Marc Benoît et Mathieu Capitaine (SAD Mirecourt), avec lesquels j'ai beaucoup appris sur le fonctionnement des exploitations agricoles et la gestion des territoires par les exploitants. Pierre-Louis Osty par son érudition tant sur les territoires des Causses que de la campagne vosgienne a su me montrer, avec enthousiasme, comment les bouleversements de l'agriculture se sont fait jour au cours des 30 dernières années. La participation aux travaux du groupe FORTE, m'a permis d'avoir une ouverture sur d'autres régions que les deux (Lorraine et Causses) sur lesquelles j'ai plus particulièrement travaillé. Merci notamment à Pascal Thinon, Elisabeth Bienaimé, Jacques Baudry, Véronique Souchère, Alain Gerbeau, Nicolas Cherman, Françoise Lemoine pour leur accueil chaleureux dans ce groupe.

Cette thèse interdisciplinaire en informatique sur une problématique agronomique a aussi donné lieu à des observations durant nos phases d'acquisition de connaissances qui ont été, sont et seront analysées par des psycho-cogniticiens et des linguistes. Ils nous donnent un nouveau regard sur ces phases si complexes et primordiales dans le cadre du développement de systèmes à base de connaissances. Je remercie particulièrement Christian Brassac et son équipe, et Lorenza Mondada pour les mises en perspectives du processus de co-construction des concepts utilisés pour décrire les connaissances du domaine de notre système.

Je tiens également à remercier tous les membres de l'équipe Orpailleur pour leur sympathie. Un merci tout particulier à Jean, Jean-Charles, Martine, Mathieu et Rim avec qui j'ai eu de longues discussions et qui n'ont cessé de m'encourager.

Merci enfin à Maude qui m'a supporté, surtout ces derniers mois, et qui a assuré des relectures fortes utiles pour éliminer une grande part des fautes d'orthographe qui sont le lot de mes écrits.

A Maude et Eliott

Résumé

Cette thèse, fruit d'une collaboration LORIA / INRA, porte sur l'étude du raisonnement à partir de cas pour aider à l'interprétation d'organisations spatiales agricoles.

Les agronomes effectuent des enquêtes auprès d'exploitants agricoles, afin d'appréhender les relations entre structure spatiale et fonctionnement des exploitations. Des modèles graphiques, appelés chorèmes d'exploitations, permettent de synthétiser ces données d'enquêtes. Sur la base de ces chorèmes, nous avons modélisé l'organisation spatiale des exploitations par des graphes bipartites aux sommets et aux arcs étiquetés par des concepts. À certains sous-graphes, décrivant des structures spatiales intéressantes, sont adjoints des explications, décrivant l'organisation fonctionnelle, pour constituer des cas. La modélisation conjointe a abouti à la reconception des chorèmes et à la constitution d'une base de cas et de hiérarchies de concepts du domaine.

Le système de RàPC ROSA permet de proposer des explications de fonctionnement à un graphe d'organisation spatiale en utilisant les cas enregistrés. La comparaison des graphes s'appuie sur le mécanisme des logiques de descriptions : étant donné un graphe d'exploitation cible, on lui associe des graphes sources de la base de cas au moyen d'opérations de généralisation et de spécialisation.

Mots-clés: Logiques de descriptions, raisonnement à partir de cas, représentation de connaissances, graphe, chorème, organisation spatiale, explication.

Abstract

The frame of this study is the case-based reasoning to bring a better understanding of relationships between spatial and functional organisation within a farm environment.

Research initial data is based on interviews of farm owners. This data was compiled in graphical model, called choremes, that highlight the core entities for analysis. Their spatial organisation, distanced from notions of metric measures or directions, enable the connection with their functional organisation. We use graphs to model these choremes. For once, graphs are useful tools for interdisciplinary communication. They also have the convenience that they can be entered directly in the computer. This effort to formalise choremes into graphs has greatly modified the appearance of choremes and brings to the constitution of a cases base and a knwoledge model domain of concepts.

Comparing graphs, with the principles of Case Based Reasoning, enable the matching of organisations. Thus the analysis of the behaviour of a farm from which we would only know the spatial description can be done by asking the system to retrieve explanation from previously analysed data.

Keywords: Description logics, case-based reasoning, knowledge representation, graph, choreme, spatial organisation, explanation.

Table des matières

Introduction **1**

1 Une recherche en informatique 1
2 Un problème d'agronomie 2
3 Un travail en inter-disciplinarité 2
4 Plan du document 3

1 La représentation dans les systèmes à base de connaissances **5**

1.1 Les premiers systèmes de représentation de connaissances 6
1.1.1 Les réseaux sémantiques 7
1.1.2 Les langages de schémas (ou frames) 8
1.2 Les approches actuelles de la représentation de connaissances 9
1.2.1 Les graphes conceptuels 10
1.2.2 La représentation de connaissances par objets 12
1.2.3 Troeps et les points de vue 14
1.3 Les logiques de descriptions 15
1.3.1 Base terminologique 16
1.3.2 Base de faits 20
1.3.3 Systèmes de logiques de descriptions 21
1.3.4 Le système Racer 22
1.4 La représentation des relations 22
1.4.1 Différents types de relations 23
1.4.2 Représentation des relations 24
1.4.3 Les relations vues comme des concepts 25
1.5 Conclusion 27

2 Raisonnement à Partir de cas **29**

2.1 Le raisonnement par analogie 30

2.1.1 Principe 30
2.1.2 Notion de similarité 31
2.2 Le raisonnement à partir de cas 32
2.2.1 Le modèle 33
2.2.2 Le cycle du RàPC 34
2.3 Le RàPC en milieu hiérarchique 36
2.3.1 L'indexation 36
2.3.2 La remémoration 37
2.3.3 La classification dure, les chemins de similarité et l'adaptation . 37
2.3.4 La classification élastique 39
2.3.5 Les difficultés de l'adaptation 41
2.4 Le raisonnement à partir d'explications 41
2.5 Conclusion 43

3 La modélisation des organisations spatiales agricoles 45
3.1 Problématique agronomique 46
3.1.1 Transformation des espaces ruraux 46
3.1.2 Production agricole et gestion de l'environnement 48
3.2 Modéliser l'organisation spatiale des exploitations agricoles 51
3.2.1 L'analyse des exploitations agricoles 51
3.2.2 La méthode des chorèmes 52
3.2.3 L'approche spatiale des exploitations agricoles 54
3.3 Exemple d'application de la méthodologie : le cas de l'étude du Sauveterre 59
3.3.1 Étape préparatoire à l'enquête 59
3.3.2 L'enquête en exploitation 60
3.3.3 Le chorème d'exploitation 64
3.4 Questions soulevées par la méthodologie appliquée 64

4 Modèle de RàPC pour les organisations spatiales agricoles 67
4.1 Le processus d'acquisition 68
4.2 Application du RàPC à la problématique agronomique 70
4.3 Des graphes pour modéliser les organisations spatiales 71
4.4 Un modèle des connaissances spatiales 72
4.4.1 Les régions 73
4.4.2 Les relations 74

4.4.3 Région ou Relation ? 77
4.5 Autres connaissances du domaine 79
4.5.1 Les explications 79
4.5.2 Les points de vue 80
4.5.3 Les échelles 81
4.6 Conclusion 81

5 Représentation des connaissances et des raisonnements 83
5.1 Sélection du langage de représentation 84
5.2 Connaissances terminologiques 84
5.2.1 Concepts 84
5.2.2 Rôles 85
5.3 Connaissances factuelles 86
5.3.1 Sommets 87
5.3.2 Graphes 88
5.3.3 Cas 90
5.4 Le RàPC dans ROSA 90
5.4.1 Principe 92
5.4.2 Classification dure 93
5.4.3 Une classification « semi-dure » 93
5.4.4 Transformation des graphes 95
5.4.5 Classification élastique 96
5.4.6 Pistes pour l'adaptation 97

6 Implantation de ROSA 99
6.1 Architecture du système ROSA 100
6.2 Modèle des données et persistance 100
6.2.1 La base terminologique 101
6.2.2 La base factuelle 102
6.3 ROSA-Interface 104
6.4 ROSA-RàPC 105
6.5 Conclusion 108

7 Intérêts et usages pour les agronomes 109
7.1 L'évolution des chorèmes 110

7.1.1 Modification d'un chorème d'exploitation agricole 111
7.1.2 Le graphe pour expliquer les relations spatiales 113
7.1.3 Le graphe pour valider la représentation chorématique 114
7.2 L'utilisation du système ROSA 115
7.2.1 La saisie d'un gos 115
7.2.2 La saisie d'un cas 116
7.2.3 Le RàPC dans ROSA 116
7.3 Usages du système 121

8 Conclusion et perspectives **123**
8.1 Réalisations 123
8.1.1 Un modèle des connaissances agronomiques 123
8.1.2 Un modèle de RàPC sur les structures spatiales 124
8.1.3 Un système utilisable 124
8.2 Enseignements 124
8.2.1 Acquisition de connaissances 124
8.2.2 Les logiques de descriptions pour la manipulation des graphes . 125
8.2.3 Tests de prototypes 125
8.3 Perpectives 126
8.3.1 Représentation et adaptation des explications 126
8.3.2 Stratégie 126
8.3.3 Rosa et les agronomes 126

Annexes **129**

A Le département Sciences de l'Action et Développement **131**
A.1 L'unité de Montpellier 132
A.2 L'unité de Mirecourt 133

B Manuel opérateur de ROSA **135**
B.1 Vue de saisie des Concepts de la base de connaissances 136
B.2 La saisie des Gos 136
B.2.1 Vue de saisie des s-entités 137
B.2.2 Vue de saisie des s-relations 139
B.2.3 Dialogue d'édition des s-entités et des s-relations 139

B.3 La vue Gos 140
B.4 La vue Cas 141
B.5 La vue RàPC 143
B.6 La vue Racer 144

Bibliographie **147**

Table des figures

1.1 Un exemple de réseau sémantique représentant la phrase « Le pré isole le champ de la rivière ». Les concepts génériques sont étiquetés en majuscule, les concepts individuels en minuscule. Les relations entre concepts sont en italique. 7

1.2 Un exemple de schéma représentant la connaissance associée au concept de SURFACE-AGRICOLE. 9

1.3 Exemple de fusion des graphes $H1$ et $H2$ en un graphe G par deux opérations de joint externe sur les sommets `pré` et `champ`. 11

1.4 Le concept de MAISON (HOUSE) d'une base de connaissance TROEPS. Figure tirée de [48]. 15

2.1 Le paradigme d'analogie. α représente la relation de ressemblance entre les univers sources (0) et les cibles (1). β représente la propriété analogique transférée. 31

2.2 Ce schéma peut se lire : sachant que `Sol(cible)` est à `Sol(source)` ce que `cible` est à `source`, quelle valeur proposer pour `Sol(cible)` étant donné `source`, `Sol(source)` et `cible` ? [72] 32

2.3 Les étapes du raisonnement à partir de cas, d'après [1]. 34

2.4 Un exemple d'adaptation à l'aide d'un chemin de similarité reposant sur une classification dure. La double flèche $\Longrightarrow$ indique les transitions entre les différents états du processus de résolution. On sait comment inverser A et C dans source, on adapte pour le faire dans cible. 38

2.5 Un exemple de chemin de similarité obtenu après classification élastique. Le bloc étiqueté par x correspond au bloc générique qui peut être substitué par tout autre bloc. 40

2.6 Un exemple d'adaptation à l'aide d'un chemin de similarité reposant sur une classification élastique. 40

3.1 Tableau des 28 chorèmes élémentaires utilisés pour la modélisation graphique des régions [11]. 53

3.2 Plaquette présentant les étapes de la modélisation d'une exploitation agricole à l'aide de la méthode d'approche spatiale de l'exploitation agricole [99]. 55

3.3 Carte du causse de Sauveterre 58

3.4 Territoire du lieu-dit de Soulages. Extrait carte 1 :25000 série bleu 2635O de l'IGN. 62
3.5 Exemple de synthèse sous forme de chorème d'une exploitation enquêtée sur le causse de Sauveterre 65

4.1 Chorème représentant l'organisation spatiale d'une exploitation agricole de la région de Vittel. Les traçages au crayon sont le résultat de l'interaction lors de la phase d'acquisition 73
4.2 Graphe d'Organisation Spatiale traduisant le chorème de la figure 4.1. . . . 74
4.3 Exemples d'explications de fonctionnement de l'exploitation agricole en relation avec son organisation spatiale décrite par le graphe de la figure 4.2. 75
4.4 Exemple de chorème représentant un siège d'exploitation au milieu de surfaces toujours en herbe avec des parcelles de céréales (carrés jaunes) plus éloignées. 75
4.5 Hiérarchie des concepts permettant de décrire les éléments du chorème. . . 76
4.6 Deux exemples de traduction de chorème en graphe. 77
4.7 Hiérarchie des concepts permettant de décrire les relations entre les entités du chorème. 78
4.8 Exemples de chorèmes rencontrés pour décrire des structures spatiales d'exploitations des causses orientées vers la production de lait de brebis. 78
4.9 Traduction sous forme de graphe pour décrire l'organisation spatiale représentée par le chorème de la figure 4.8a. 78
4.10 Exemple d'adaptation d'une explication. 80
4.11 Exemple de cas pour lequel un point de vue d'attrait des brebis est présent. 80

5.1 Hiérarchie des concepts manipulés dans ROSA. 85
5.2 Hiérarchie des rôles $\mathcal{H}_R$ utilisés pour la représentation des connaissances dans notre système. 86
5.3 Exemple de graphe décrivant des connaissances sur une exploitation. 87
5.4 Hiérarchie partielle des index de cas (les explications ne sont pas représentées). 92
5.5 Exemple de remémoration à l'aide de la classification dure des graphes. . . 94
5.6 Graphe des relations de voisinage pour les relations de la théorie RCC-8 [24]. 95
5.7 Deux exemples de règles de transformation spatiale. 96
5.8 Exemple de remémoration à l'aide de la classification élastique des graphes. 97

6.1 Architecture du système de Raisonnement sur les Organisations Spatiales Agricoles (ROSA). 101
6.2 Un exemple de saisie ou de modification des entités d'un chorème dans ROSA-Interface. 106
6.3 Difficulté d'appariement de graphe dans le cadre des LD ; les sommets S'_1 et S'_4 doivent représenter un seul et même individu, ce que ne permet pas de garantir le système de LD. 107

7.1 Chorème représentant l'organisation spatiale d'une exploitation agricole de l'est de la France. 111

7.2 Nouveau chorème représentant la même exploitation que celle modélisée figure 7.1. 112
7.3 Composition chorématique finale de l'exploitation présentée figure 7.1 après la phase d'acquisition. 114
7.4 Exemple d'un nouveau chorème étudié. 117
7.5 Graphe d'organisation d'une nouvelle exploitation que l'on cherche à analyser. 117
7.6 Classification dure du Gos de la figure 7.5 sur la base de cas. Une partie du `gos(cible)` est appariée à un cas de la base de cas. 118
7.7 Les cinq nouveaux cas appariés à l'aide de la classification "semi-dure" avec leur indice de similarité. 119

B.1 Présentation de l'application à son ouverture. 135
B.2 Edition des concepts, rôles et points de vue. 137
B.3 La vue s-entité met en regard la liste des entités spatiales et le dessin du chorème d'exploitation. 138
B.4 Fenêtre de dialogue pour saisir les informations sur l'entité éditée. 139
B.5 Dialogue pour sélectionner un s-entité relié au s-relation. 140
B.6 Présentation d'un `gos` avec la sélection d'un `gos-e` mis en relief. 141
B.7 Visualisation et édition des cas. 142
B.8 Visualisation des appariements pour les cas remémorés. 143
B.9 Cette vue permet l'accès direct aux fonctions de RACER. 144

Introduction

L'objet de cette thèse est l'étude et le développement d'un système de raisonnement fondé sur le modèle du raisonnement à partir de cas (RÀPC). Ce formalisme s'inspire d'un mode de raisonnement communément utilisé par l'être humain, à savoir le raisonnement par analogie. Ce raisonnement s'appuie sur des unités d'expériences passées pour résoudre un problème posé. Pour ce raisonnement, on définit des cas, les unités d'expérience, constituées de deux parties. La première décrit un problème. La deuxième représente la solution adoptée. Le raisonnement se déroule en deux grandes étapes : recherche d'un ou plusieurs cas dont les parties problèmes sont similaires au nouveau problème posé, puis adaptation des solutions enregistrées afin de proposer une solution au problème présent.

1 Une recherche en informatique

Cette thèse se situe dans la branche de l'intelligence artificielle qui traite de représentation et de manipulation des connaissances. La question que soulève cette recherche est : « comment décrire et manipuler des connaissances pour construire un système à base de connaissances ? » Pour mener à bien notre recherche, deux champs d'étude sont à considérer :

- La représentation de connaissances et notamment pour ce qui nous concerne le problème de la représentation des relations spatiales,
- Le raisonnement à partir de cas pour la manipulation des connaissances représentées.

Notre travail s'inscrit dans la lignée de travaux antérieurs effectués dans l'équipe de recherche Orpailleur au LORIA[1]. Du côté de la représentation, la question du choix du formalisme le mieux adapté a été posée, et nous nous intéressons, dans le cadre de cette thèse, aux capacités des logiques de descriptions pour représenter des structures complexes et des relations spatiales. Pour le RÀPC, nous nous plaçons dans la poursuite des travaux de Jean Lieber.

Le développement d'un système à base de connaissances implique un travail d'acquisition des connaissances important. Les langages informatiques étant moins expressifs que le langage naturel, il y a une certaine difficulté à exprimer puis à représenter la connaissance dans un système informatique. Nous avons appliqué une démarche de modélisation de représentation et de raisonnement sur des données issues de l'agronomie. Ce travail expérimental s'est déroulé sous forme d'entretiens avec des experts du domaine sous le regard de chercheurs en psychologie cognitive, spécialistes du travail collaboratif, et de

[1]Laboratoire lorrain de recherche en informatique et ses applications

linguistes.

2 Un problème d'agronomie

Le problème posé par les chercheurs du département des Sciences de l'Action et Développement de l'INRA[2] (INRA-SAD) fait suite à vingt cinq ans d'observations des transformations sociales et spatio-environnementales. Sur le causse Méjan, les gestionnaires du Parc National des Cévennes et la profession agricole se sont intéressés aux conditions du maintien des paysages ouverts par les élevages et se sont appuyés sur l'analyse de la dynamique des systèmes techniques étudiés par les agronomes. En Lorraine, la société des Eaux de Vittel, soucieuse de conserver la qualité et la pureté de l'eau, s'est fortement impliquée pour développer une réflexion afin d'assurer une gestion de l'espace agricole qui soit en rapport avec ses préoccupations. Elle s'est tournée vers les agronomes du SAD afin qu'ils pilotent les transformations nécessaires des systèmes de production dans cette région de l'est de la France.

Ces deux séries d'observations, différentes dans leurs formes, ont suggéré aux agronomes une hypothèse à tester : "Il existe un lien étroit entre l'organisation spatiale et le fonctionnement d'une exploitation agricole". Autrement dit : "Les agriculteurs utilisent au mieux l'espace en fonction de ses contraintes propres et des contraintes liées à l'environnement socio-économique. En outre, dans la mesure du possible, ils aménagent et restructurent leur espace (on parle d'espace produit) afin de transformer les contraintes en atouts : ce sont les opérations de configuration de l'espace". Encore faut-il comprendre comment ces contraintes sont prises en compte ; comment sont liés organisation spatiale et fonctionnement.

3 Un travail en inter-disciplinarité

Dans le souci de tester leurs hypothèses, de les instrumenter, ces agronomes ont donc cherché à formaliser leurs intuitions et leur observations afin de leur permettre de comparer des situations apparemment différentes.

Pour cela, ils ont développé un outil de formalisation et de représentation graphique, initialement conçu par les géographes : les chorèmes, qu'ils ont adaptés à leurs besoins. Ils se sont également tournés vers les informaticiens, chercheurs en intelligence artificielle, afin de développer les méthodes et outils à même de comparer différentes situations et déterminer en détail les liens entre organisation spatiale et fonctionnement des exploitations. Dans ce cadre, nous avons proposé l'utilisation des graphes pour traduire les organisations spatiales des exploitations associées à des explications faisant le lien entre une organisation spatiale observée et le fonctionnement de l'exploitation qui en résulte.

Partant des connaissances détaillées acquises au cours de leur travail, nous avons proposé le développement d'un système de raisonnement à partir de cas ou plus exactement de raisonnement à partir d'explications. Ce système a pour but de proposer des explications

[2]Institut National de Recherche Agronomique

sur le fonctionnement d'une exploitation à partir de la description de son organisation spatiale en utilisant les connaissances acquises précédemment sur d'autres exploitations.

Cette thèse fait suite à plusieurs autres collaborations entre agronomes et informaticiens. L'utilisation des chorèmes est une des dernières étapes mises en place par les agronomes du SAD, mais les problématiques d'étude et de raisonnement sur les organisations spatiales sont des thématiques anciennes.

Cette collaboration a démarré entre l'INRA et le LORIA dans le cadre de la thèse de F. Le Ber [61] qui portait sur l'interprétation d'images satellitaires. Elle s'est poursuivie par la thèse de L. Mangelinck [79, 80] qui a développé un système de reconnaissance de structures spatiales des territoires villageois également sur des images satellitaires.

Au delà de la reconnaissance et de la classification de structures spatiales, on voudrait pouvoir interpréter ces structures, comprendre leurs origines et leurs évolutions. Cette étape supplémentaire est étudiée dans le cadre de cette thèse. L'interprétation consiste ici à déterminer en quoi l'organisation spatiale peut être un révélateur du fonctionnement d'une exploitation. On ne cherche pas à reconnaître mais à interpréter (même s'il faut d'abord reconnaître).

Les agronomes cherchent donc à obtenir un système offrant une automatisation des raisonnements associant une explication du fonctionnement d'une exploitation en fonction de son organisation spatiale.

La réalisation d'un tel système doit permettre de formaliser plus avant les connaissances que possèdent les agronomes sur les exploitations agricoles et sur l'utilisation des chorèmes. En effet, les questions posées sur l'utilisation des chorèmes sont nombreuses. Nous en citons ici trois parmi les plus importantes :

- Les chorèmes sont-ils suffisamment génériques ? Les connaissances acquises sur les exploitations d'une région donnée peuvent-elle être utilisées pour expliquer le fonctionnement d'exploitations dans des régions différentes ?
- Les chorèmes sont-ils reproductibles ? Deux agronomes, avec des expériences et des connaissances sur des territoires différents produiront-ils les mêmes chorèmes ?

Notre projet est de proposer aux agronomes un outil susceptible de les aider à répondre à ces questions. Il n'a pas pour but d'y répondre directement. En revanche, le travail de modélisation effectué en collaboration avec les agronomes a pour effet collatéral de revisiter leurs connaissances et leur façon de faire (chapitre 7). Cet effet a été mis en évidence lors de deux expériences particulières, menée sous le regard et l'analyse des psychologues de la cognition et de linguistes ethnométhodologues qui ont interrogé les processus de modélisation tant du coté des agronomes que du coté des informaticiens [10, 96].

4 Plan du document

La construction du système de raisonnement à partir de cas que nous étudions ne peut s'envisager sans l'utilisation d'un langage de représentation de connaissances adapté. Le chapitre 1 est consacré à l'étude de ces langages et des mécanismes de raisonnement associés. L'histoire des systèmes de représentation de connaissances commence avec les réseaux sémantiques et les langages de "frames". Ces deux voies donneront naissance à

de nombreux autres systèmes tels que les systèmes à base de graphes conceptuels, de représentation de connaissances par objets ou encore des logiques de descriptions. Nous nous attacherons à décrire cette dernière famille de langages et plus particulièrement le langage du système RACER que nous avons choisi comme système de représentation des connaissances dans le cadre de notre étude.

Le chapitre 2 replace le raisonnement par analogie dans le contexte des recherches en intelligence artificielle. Le principe du raisonnement à partir de cas est détaillé et son application en milieu hiérarchique est envisagée.

Le chapitre 3 présente le contexte agronomique dans lequel sont réalisées les études des agronomes sur l'organisation spatiale des exploitations agricoles. Ces études découlent de l'évolution récente de l'espace rural. La maîtrise de cette évolution et des contraintes qu'elle impose passe par l'étude de l'impact sur le territoire de l'activité agricole qui est l'utilisatrice principale de l'espace. La méthode de modélisation de l'organisation spatiale des exploitations agricoles est décrite. Son utilisation est présentée sur un cas concret d'étude menée sur des exploitations sur le causse de Sauveterre où les agronomes cherchent à identifier la réponse des exploitants à l'embroussaillement du territoire. Comment comparer les réponses des exploitants en prenant en compte les contraintes particulières du territoire de chaque exploitation ? Comment garantir la validité et la reproductibilité des résultats ? La réponse à ces questions conditionne la qualité des analyses faites et l'utilisation de telles méthodes d'études des exploitations dans d'autres contextes. La mise en place d'un partenariat pluridisciplinaire avec des chercheurs en intelligence artificielle doit permettre d'apporter des réponses à ces questions.

Le chapitre 4 expose la proposition de développement d'un système de RÀPC en milieu hiérarchique pour répondre aux besoins des agronomes. L'acquisition et la modélisation des connaissances sont explicitées.

Au chapitre 5, nous décrivons le modèle des données implanté en logique de descriptions. Puis, nous présentons le modèle de remémoration et ses différents modes d'application. Enfin, nous donnons des pistes pour le développement de l'adaptation.

Le chapitre 6 présente l'implantation des modèles de données et de raisonnement au sein de notre système de Raisonnement sur les Organisations Spatiales Agricoles (ROSA). Nous y révélons les difficultés rencontrées dans la manipulation de structures complexes représentées par les graphes en logique de descriptions . Nous présentons comment nous avons apporté une solution au problème à l'aide d'une implantation de la remémoration en s'appuyant sur les logiques de descriptions et les fonctionnalités du langage de programmation JAVA.

Le chapitre 7 couvre la réflexion sur les chorèmes liée à leur transformation en graphe. Les résultats de l'utilisation du système ROSA sont abordés du point de vue agronomique et le point est fait sur les forces et faiblesses d'un tel système en proposant des voies d'amélioration pour rendre ce système plus efficace.

La conclusion présente l'état actuel du système. De nombreux travaux restent à mener afin d'avoir un système réellement fonctionnel, mais les résultats déjà obtenus sont prometteurs. Enfin nous ouvrons les perspectives sur l'utilisation d'un tel système dans d'autres contextes.

Chapitre 1

La représentation dans les systèmes à base de connaissances

Sommaire

1.1 Les premiers systèmes de représentation de connaissances 6
1.1.1 Les réseaux sémantiques 7
1.1.2 Les langages de schémas (ou frames) 8
1.2 Les approches actuelles de la représentation de connaissances 9
1.2.1 Les graphes conceptuels 10
1.2.2 La représentation de connaissances par objets 12
1.2.3 Troeps et les points de vue 14
1.3 Les logiques de descriptions 15
1.3.1 Base terminologique 16
1.3.2 Base de faits 20
1.3.3 Systèmes de logiques de descriptions 21
1.3.4 Le système Racer 22
1.4 La représentation des relations 22
1.4.1 Différents types de relations 23
1.4.2 Représentation des relations 24
1.4.3 Les relations vues comme des concepts 25
1.5 Conclusion 27

« La connaissance est le rapport entre un sujet (connaissant) et un objet (connu). Toute connaissance repose sur un langage qui conditionne le rapport au monde visé par la connaissance » [3].

Les langages sont de deux ordres : naturels ou artificiels. Les langues naturelles sont les langues parlées comme le français, l'anglais etc. Les langues artificielles sont mathématiques ou logiques. De ces deux formes de langages, on déduit deux formes de connaissances : scientifiques ou phénoménologiques suivant que les langages utilisés pour les décrire sont artificiels ou naturels.

Les langues artificielles sont calculatoires et formelles. Les expressions dans ces langages sont manipulées en fonction de leur forme, ou contenu syntaxique et non en fonction de leur contenu sémantique. Au contraire, les langues naturelles ne sont pas calculatoires : les expressions sont manipulées en fonction de leur contenu sémantique et non en fonction de leur syntaxe.

La langue naturelle est d'abord un contenu sémantique que l'on organise par une syntaxe alors que la langue scientifique est une syntaxe à laquelle on associe une sémantique. Dans la première, dire c'est signifier. Dans la seconde, dire c'est calculer [3].

Une des questions de l'intelligence artificielle (I.A.) est de savoir dans quelle mesure il est possible d'étudier la sémantique à travers un niveau syntaxique. C'est la question à laquelle les langages de représentation de connaissances tentent de répondre.

Dans ce chapitre, nous présentons les premiers travaux sur les langages de représentation de connaissances : *réseaux sémantiques* et *frames*, avec les raisonnements qui leur sont associés et en présentant leurs limites. Dans une seconde section, nous montrons comment, sur la base de ces travaux initiaux, se sont développés de nouveaux langages tels que les *Graphes Conceptuels* et les *Systèmes de Représentation de Connaissances par objets*. Nous présentons plus largement, dans la section 1.3, les *logiques de descriptions*, que nous avons utilisées dans notre système. La section 1.4 est consacrée au problème particulier de la représentation des relations dans les différents systèmes envisagés.

1.1 Les premiers systèmes de représentation de connaissances

La représentation de connaissances est un aspect fondamental du développement des systèmes en intelligence artificielle. Elle doit offrir une représentation syntaxique et des opérateurs assurant une correspondance entre dénotation syntaxique et interprétation sémantique [3]. Elle doit être combinée avec des techniques de raisonnement pour inférer de nouvelles connaissances. La représentation permet d'introduire une communication homme-machine aisée. Elle sera d'autant plus intéressante qu'elle offrira à l'utilisateur de suivre ou d'intervenir sur le déroulement des raisonnements.

Les premières formes de représentation de connaissances étaient basées sur la logique classique avec des raisonnements par inférences logiques monotones. Puis très vite, des systèmes à base de règles de la forme "Si...alors" ont vu le jour, mettant en place des mécanismes de chaînage avant et arrière pour simuler un raisonnement de type cause-effet. Ces formalismes décrivent les connaissances sous forme de proposition ou affirmation sur

le monde.

Historiquement, un des premiers langages génériques s'appuie sur le modèle des *réseaux sémantiques* mis en œuvre par Quillian, qui propose une représentation de connaissances prenant en compte la structure et les relations entre les propositions décrivant le monde [110]. C'est un modèle de la mémoire associative humaine qui se caractérise par une organisation de la connaissance sous forme de graphes orientés et étiquetés. Il fut suivi par les *langages de schémas (ou frames)* construits autour de la notion de schéma : un nom auquel est associé un ensemble d'attributs [95].

1.1.1 Les réseaux sémantiques

Dans les réseaux sémantiques, la connaissance est représentée par des graphes étiquetés constitués de nœuds et d'arcs. Les nœuds sont étiquetés par des concepts génériques ou individuels. Les arcs, étiquetés par des relations de différents ordres, décrivent des liens entre concepts. Par exemple, le réseau présenté figure 1.1 représente la phrase « Le pré isole le champ de la rivière ». Des opérations de manipulation de graphes permettent de raisonner sur un réseau : création ou suppression de nœuds ou d'arcs, parcours du graphe des relations, inférence par héritage et filtrage.

Les relations *sorte-de* (entre concepts génériques) et *est-un* (entre un concept individuel et un concept générique) ont été différenciées des autres relations pour traduire un lien de généralité entre les concepts. Ces deux relations sont souvent confondues sous le seul terme *sorte-de* et l'inférence par héritage peut se ramener à un parcours du graphe de cette relation. Dans l'exemple précédent, le « pré contient de l'herbe » peut être inféré par la relation de spécialisation `pré` en SURFACE-TOUJOURS-EN-HERBE.

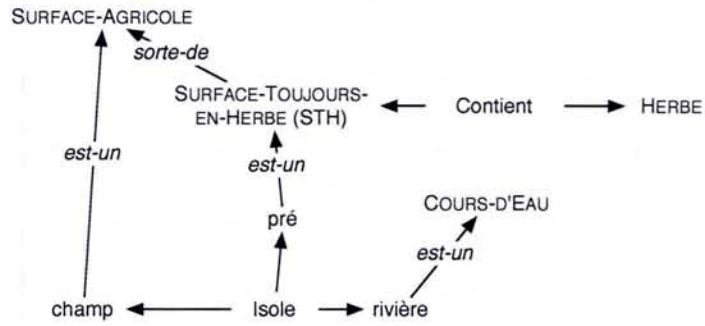


FIG. 1.1 – Un exemple de réseau sémantique représentant la phrase « Le pré isole le champ de la rivière ». Les concepts génériques sont étiquetés en majuscule, les concepts individuels en minuscule. Les relations entre concepts sont en italique.

Les réseaux sémantiques, tout en restant propositionnels comme la logique et les systèmes à base de règles, reconnaissent déjà l'importance de la structure de la connaissance à travers les nœuds et les arcs des graphes. Cette représentation rend visibles les diverses relations existantes entre les objets, ainsi que la notion de "distance" entre deux concepts (nombre d'arcs du chemin connectant les deux concepts).

Mais les réseaux sémantiques souffrent d'un manque de sémantique ! Le sens donné aux relations dépend de l'application et n'est en général pas reconnu dans les systèmes implantant des réseaux sémantiques. Le problème de la granularité, du choix des nœuds et des liens de base n'est pas évident. Quel est le niveau de détail adéquat et quelles structures générales sont nécessaires à l'expression et à la compréhension d'une proposition ? Par ailleurs, la seule information attachée à un nœud est son nom donné par une étiquette. Il est difficile d'exprimer des propositions ayant des quantificateurs universels, existentiels et numériques. Par exemple, le sous-graphe : SURFACE TOUJOURS EN HERBE ← *Contient* → HERBE, signifie-t-il que « toutes les surfaces toujours en herbe contiennent de l'herbe », ou que « quelques surfaces toujours en herbe contiennent de l'herbe » ?

1.1.2 Les langages de schémas (ou frames)

Les langages de schémas désignent les formalismes de représentation qui sont issus, en grande partie, de la théorie des *frames* de Marvin Minsky [95]. Le principe de ces langages est qu'un concept est représenté par un individu particulier que l'on nomme *stéréotype*. Tout stéréotype possède une description. Cette description forme un *frame* décrivant un concept avec ses différentes propriétés ou attributs ("slots"). Les attributs décrivent un point de vue sur l'entité et sont eux-mêmes décrits par des facettes qui sont de deux types :

- Déclaratives : elles définissent des contraintes pour la valeur de l'attribut. Ces contraintes portent sur son type, son co-domaine (valeurs possibles) ou la valeur sûre (toujours vraie) ou la valeur par défaut (la plus probable) pour cet attribut ;
- Procédurales : elles introduisent des procédures, appelées réflexes, qui peuvent être vues comme la représentation locale de procédures de raisonnement. Les réflexes peuvent être déclenchés lors de l'accès à un attribut en lecture ou en écriture. En particulier, lorsque l'attribut n'a pas de valeur courante, celle-ci peut être calculée ou positionnée à la valeur par défaut.

Les schémas sont organisés par niveau de généralité, en une hiérarchie d'héritage. Un schéma est une spécialisation d'un ou plusieurs super-schémas. L'exemple de schéma donné figure 1.2 décrit le concept de SURFACE-AGRICOLE défini par un ensemble d'attributs qui peuvent avoir plusieurs facettes.

Cette représentation s'appuie sur le paradigme selon lequel, pour analyser une situation nouvelle, un être humain utilise des situations particulières auxquelles il a déjà été confronté. Les caractéristiques de la situation nouvelle sont utilisées afin de rechercher dans la base de connaissances tous les frames qui peuvent être associés à cette situation. La rapidité des activités mentales humaines s'explique, dans le cadre de la théorie des prototypes, par le fait qu'une situation nouvelle est identifiée dans la mémoire avec la situation prototypique la plus appropriée. Les informations générales de ce prototype sont ensuite récupérées et éventuellement modifiées pour créer la représentation précise de la nouvelle situation.

En fait, deux approches des schémas existent. Celle décrite ci-avant est la plus fidèle aux idées de Minsky, mais elle n'est pas la plus utilisée. Par opposition, l'approche définitionnelle distingue deux types de schémas : les schémas de classes et les schémas d'instances. Un individu est classé dans la classe la plus proche du prototype retrouvé.

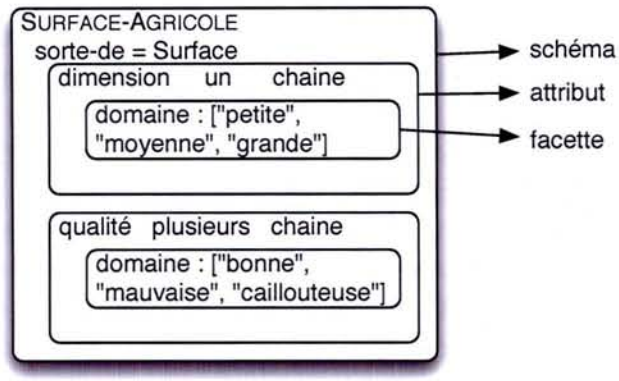


FIG. 1.2 – Un exemple de schéma représentant la connaissance associée au concept de SURFACE-AGRICOLE.

Cette mise en correspondance est effectuée par une opération nommée *filtrage* qui consiste à sélectionner un ensemble de schémas présentant des caractéristiques communes avec l'individu à classer. Un nouveau schéma est positionné dans la hiérarchie des schémas par un mécanisme de classification qui permet de trouver sa place et de l'insérer dans la hiérarchie en mettant à jour, au besoin, les liens de spécialisation des schémas existants.

Les schémas regroupent les connaissances d'une situation, d'un événement ou d'un objet dans des paquets (“*chunks*”) accessibles en tant qu'unités. Ils définissent une organisation relativement structurée de la connaissance contrairement à la forme propositionnelle des réseaux sémantiques, de la logique ou des systèmes à bases de règles. Ils permettent la modélisation dans des domaines où les objets manipulés sont complexes et riches. Le regroupement des informations sous forme de paquets offre une représentation structurelle concise et facilement exploitable.

Mais la plupart des systèmes à base de schémas manquent d'un formalisme de base qui induit un manque de rigueur dans l'utilisation des différentes notions. Par exemple, l'attachement procédural pour un calcul est très souvent utilisé lors de la classification et la facette défaut, qui représente une valeur incertaine mais probable, est en général utilisée pour faire des inférences certaines dans un contexte monotone. Par ailleurs, les systèmes de schémas ne font pas de distinction entre les caractéristiques définitionnelles d'un schéma et les caractéristiques déduites [81].

1.2 Les approches actuelles de la représentation de connaissances

La rationalisation de ces systèmes a conduit à l'énoncé de principes donnés par R. Brachman [9] :

- Ne pas considérer la même entité, tantôt comme un individu (une voiture rouge particulière) et tantôt comme un ensemble d'individus (la classe des voitures rouges) ;
- Ne pas confondre propriétés descriptives (les voitures ont des roues) et propriétés

typiques (les voiture roulent) ;
- Ne pas confondre définition (conditions nécessaires et suffisantes d'appartenance à une classe), description (conditions nécessaires) et prototype (valeurs typiques des attributs).

La formalisation des réseaux sémantiques a conduit J.-F. Sowa à proposer le formalisme des graphes conceptuels. Le développement de la notion d'objets, initialement envisagée dans le cadre du génie logiciel, a abouti au développement de systèmes de représentation de connaissances par objets. Ces derniers sont aussi fortement inspirés des langages de schémas. Enfin, les travaux de rationalisation de la représentation de connaissances, tant des langages de schémas que des réseaux sémantiques, ont amené à la définition du formalisme des *logiques de descriptions* (LD).

1.2.1 Les graphes conceptuels

Les *graphes conceptuels* ont été introduits par Sowa [120]. Celui-ci présente ces graphes comme un intermédiaire entre la langue naturelle et la logique des prédicats. Le formalisme est inspiré des *réseaux sémantiques* et se caractérise par une organisation des connaissances en un graphe orienté.

Un graphe conceptuel est un graphe fini, bipartite, étiqueté, non nécessairement connexe [97]. Les deux classes de sommets sont étiquetées respectivement par des noms de "concepts" et des noms de "relations conceptuelles" entre ces concepts. Les types des concepts sont organisés dans une hiérarchie induite par la relation de *spécialisation* : (CHAMP spécialise SURFACE-AGRICOLE). Les relations conceptuelles sont définies par un profil désignant le type des concepts reliés par la relation. L'ensemble des types de relations est ordonné par une relation de spécialisation. Cependant, deux types de relations d'arités différentes sont considérées comme incomparables.

Les sommets concepts d'un graphe peuvent être génériques ou individuels. Les concepts génériques correspondent aux variables en logique et représentent des individus non spécifiés. Les concepts individuels, identifiés par des marqueurs, correspondent aux constantes de la logique et décrivent des individus particuliers du monde.

Le vocabulaire conceptuel du modèle des graphes conceptuels est principalement composé de trois ensembles disjoints :
- L'ensemble T_c des types de concepts ;
- L'ensemble T_r des types de relations ;
- L'ensemble M des marqueurs individuels.

Ces trois ensembles sont les composants essentiels du *support* sur la base duquel les graphes conceptuels peuvent être manipulés et interprétés.

Pour contrôler la sémantique des graphes conceptuels, Sowa propose de définir un système formel. Un ensemble initial de graphes, appelé *base canonique*, contient des graphes déclarés sémantiquement corrects. De nouveaux graphes canoniques peuvent être dérivés des graphes canoniques déjà présents dans la base à l'aide des règles de formation suivantes [69] :
- Restriction de concept. Le type d'un sommet concept générique peut être remplacé par un type plus spécifique ou par un marqueur individuel de même type.

- Joint interne. Si dans un graphe, deux sommets possèdent un même marqueur individuel, alors ces sommets sont fusionnés en un seul.
- Joint externe. Deux graphes possédant chacun un sommet concept de même type peuvent être joints en un seul graphe par fusion des deux sommets concepts.

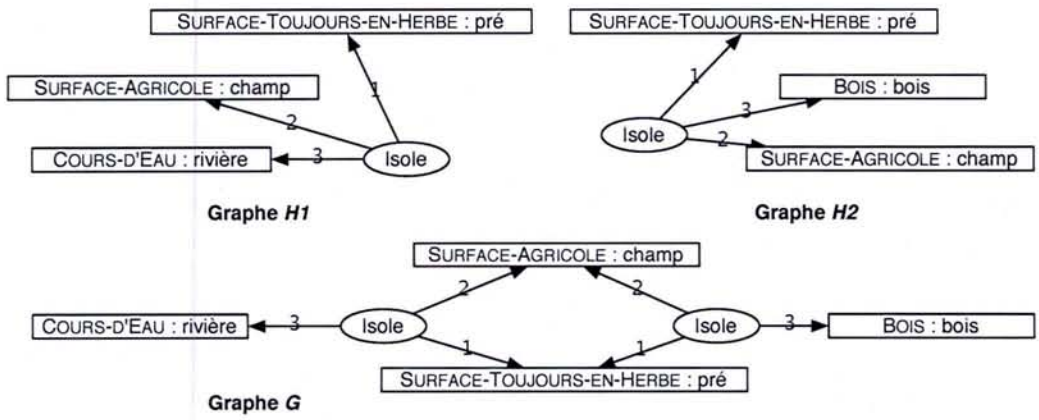


FIG. 1.3 – Exemple de fusion des graphes $H1$ et $H2$ en un graphe G par deux opérations de joint externe sur les sommets `pré` et `champ`.

Le graphe conceptuel $H1$ de la figure 1.3 représente la phrase « Le pré isole le champ de la rivière ». De même le graphe $H2$ décrit la situation « Le pré isole le champ du bois ». Ils peuvent être fusionnés en un graphe G par deux opérations de joint externe. Les deux sommets `pré` (resp. `champ`) des graphes $H1$ et $H2$ sont fusionnés et connectent ainsi les deux graphes $H1$ et $H2$ en un graphe G par les sommets résultants `pré` (resp. `champ`).

Sowa relie la définition sous forme de graphe à une définition en logique en associant à chaque graphe une formule logique $\Phi(\mathrm{G})$. Une *relation de spécialisation* entre graphes conceptuels correspond à l'existence d'un morphisme entre les graphes et est définie mathématiquement par la formule [20] :

$$G \leq H \text{ si et seulement si } \Phi(G) \vdash \Phi(F).$$

L'opération de projection est l'opération fondamentale du modèle des graphes conceptuels. C'est un morphisme de graphes étiquetés, qui permet le calcul effectif de la relation de spécialisation sur les graphes, en autorisant une restriction des étiquettes des sommets si cette restriction respecte les relations d'ordre du support.

Les graphes conceptuels peuvent être considérés comme un modèle *déclaratif* de représentation de connaissances basé sur des graphes étiquetés ayant une *sémantique logique consistante et complète.* Les calculs d'inférence peuvent également être réalisés par des *algorithmes de manipulation de graphes* sans utiliser un démonstrateur de théorèmes en logique.

L'intérêt des graphes conceptuels tient également dans leur représentation graphique qui facilite la présentation des connaissances manipulées par un utilisateur tout en offrant un formalisme rigoureux. Contrairement aux réseaux sémantiques, les graphes conceptuels organisent la connaissance dans différentes structures constitutives du support. Cette

structuration entraîne une plus grande clarté lors de l'utilisation de ce modèle pour l'acquisition ou la représentation de connaissances. Cette distinction permet la manipulation et la vérification de chaque structure à partir d'un ensemble réduit de règles de cohérence.

1.2.2 La représentation de connaissances par objets

Les langages à objets ne sont pas à proprement parler des langages de représentation de connaissances mais des langages de programmation. Toutefois, leurs mécanismes se retrouvent dans beaucoup de systèmes de représentation de connaissances. Le succès des "objets informatiques" est justifié habituellement par les qualités que recherche et promeut le génie logiciel : modularité, extensibilité ou réutilisabilité sont des références obligées. Pourtant, une autre raison est sans doute aussi importante : les "objets informatiques" ont une capacité naturelle de représentation des "objets du monde réel". Cette capacité est non seulement à l'origine de la représentation des connaissances par objets en intelligence artificielle, mais aussi des méthodes d'analyse et de conception en génie logiciel ou encore des bases de données à objets [102]. Ces langages permettent la manipulation d'objets pour lesquels on peut définir des propriétés à l'aide d'attributs et de méthodes. Les mots centraux dans les langages à objets sont : *classe et instance.*

- La classe est une structure (un type) à partir de laquelle s'engendrent des instances.
- Une classe définit, encapsule, les attributs et les comportements de ses instances.
- Les classes sont organisées en une hiérarchie. La relation d'héritage autorise la redéfinition ou l'ajout de propriétés (attributs et méthodes) des classes par rapport à leur classe mère. Beaucoup de systèmes d'objets ne définissent que l'héritage simple. Dans ce cas, le graphe de la hiérarchie est un arbre.
- Les attributs dans une classe `C` peuvent être vus comme des couples (type, valeur) où type indique le type des valeurs qui seront associées aux attributs des instances de la classe `C`. Un type peut être primitif (chaîne de caractères, entier, flottant, etc.) ou donné par une classe `D`, auquel cas l'attribut définit une relation entre `C` et `D` (`C` est le domaine de la relation et `D` est le co-domaine de la relation). Les attributs typés par une classe permettent de définir des liens vers d'autres instances assurant ainsi la possibilité des mécanismes d'envoi de messages.
- Le comportement des instances est modélisé dans des méthodes qui sont exécutées par le mécanisme d'envoi de messages entre instances.
- Les instances sont créées à partir de constructeurs définis dans les classes qui permettent d'affecter leur valeur à chacun des attributs.

Ainsi, parallèlement au développement de l'approche objet en génie logiciel, sont apparus des langages hybrides de représentation inspirés des langages de programmation par objets mais aussi des langages de *frames* et des réseaux sémantiques. C'est à partir des langages hybrides de représentation que s'est développée la notion de *représentation de connaissances par objets* (RCO). Dans le même temps, des recherches plus fondamentales ont été poursuivies. Des travaux ont conduit à évaluer la décidabilité et la complexité de certains mécanismes de déduction. Des recherches se sont alors principalement focalisées sur un type de langages nommés *logiques de descriptions* (voir 1.3).

Les langages de RCO représentent un compromis entre les langages de programmation par objets, auxquels ils empruntent les concepts et les fonctionnalités procédurales, et les

logiques de descriptions pour leurs aspects sémantiques. La notion de RCO se caractérise plus par la disponibilité et l'application de systèmes que par un corpus théorique uniforme. Mais la recherche théorique sur la RCO évolue vers une formalisation croissante qui conduit à définir une sémantique dénotationnelle des systèmes de RCO, à évaluer la décidabilité et la complexité des mécanismes d'inférences disponibles, et à concevoir des algorithmes corrects pour ces mécanismes.

Cependant, l'effort de formalisation ayant poussé à une simplification des systèmes existants, de nombreuses opérations et innovations - souvent très utiles sur le plan pratique - sont laissées dans l'ombre par le travail de formalisation. Même les systèmes développés sur la base des travaux théoriques offrent des traits qui sortent de la théorie, mais qui les rendent utiles. Il est donc important de s'intéresser à tous les aspects d'un système de RCO et non uniquement à ceux qui sont dûment formalisés [102].

À proprement parler, il n'y a pas encore aujourd'hui de principes intangibles sur lesquels reposent les systèmes de RCO. Il est cependant possible de décrire à gros traits ce qu'est un système typique. La fonction d'un tel système est d'organiser les connaissances autour de la notion d'objet et de fournir des services inférentiels destinés à compléter l'information disponible. Ces services sont notamment la classification et les facettes d'inférence de valeurs d'attributs.

La classification dans les RCO peut prendre deux formes :

- La *classification de classe* consiste à placer une nouvelle classe dans le graphe des classes en recherchant ses super-classes les plus spécifiques et ses sous-classes les plus générales. Le test de spécialisation se fait sur la base de la structure des classes. La sémantique de la relation de spécialisation est celle de l'inclusion ensembliste qui permet de définir, entre autres, les principes suivants :
  - Plus une classe est spécifique, plus les domaines de ses attributs sont restreints ;
  - Si une classe est sous-classe de plusieurs autres, les domaines de ses attributs sont inclus dans (ou égaux à) l'intersection des domaines des attributs des super-classes ;
  - Si un objet est instance d'une classe, il l'est aussi de toutes les super-classes.
- La *classification d'instance* qui permet d'identifier la classe d'un individu.

Les facettes d'inférence de valeurs d'attributs que l'on trouve classiquement sont :

- Valeur par défaut, qui permet de fixer une valeur donnée en l'absence d'autres valeurs ;
- Passage de valeur, qui contraint la valeur à être celle d'un autre attribut ;
- Filtrage, qui donne pour valeur d'un attribut l'ensemble des objets qui satisfont certaines contraintes ;
- Attachement procédural, qui provoque l'exécution d'une fonction retournant la valeur demandée.

Plusieurs systèmes de RCO ont été développés. Quelques représentants français de ces systèmes sont : SHIRKA et son successeur TROEPS [48], YAFOOL [31] et le dernier né AROM [109]. En particulier, YAFOOL *Yet Another Frame-based Object Oriented Language*, dont nous parlerons plus en détail au paragraphe 1.4.3 a été utilisé dans des travaux proches des nôtres : pour un système de raisonnement à partir de cas en chimie (RESYN/RÀPC) [72] et un système de reconnaissance de structures spatiales [79].

1.2.3 Troeps et les points de vue

TROEPS permet de faire de la classification d'instances dans un modèle de connaissances multi-points de vue. Chaque instance fait partie d'un concept défini. Elle est ensuite classée, suivant différents points de vue, dans les hiérarchies de classes décrivant ces points de vue.

Chaque concept possède des attributs clefs, qui servent à identifier les instances, et d'autres attributs. Un attribut est constitué d'un nom, d'un type choisi parmi les types de base (entier, chaîne, booléen ou réel) ou les autres concepts. La cardinalité d'un attribut est soit unique, soit liste (une instance peut apparaître plusieurs fois dans une liste), soit un ensemble. Un attribut représente une propriété, un composant, ou encore un lien vers une (ou des) instance(s) d'un autre concept. Les relations de composition possèdent le sens de la composition physique (composant / objet) au sens où une table est composée de quatre pieds [Winston, 87]. Ainsi TROEPS assure pour le champ de type composant que :

- Le composant existe ;
- Qu'il n'y a pas de cycle i.e. qu'un objet ne peut pas être un composant d'un de ses composants ;
- Qu'il n'y a pas partage d'un composant dans plusieurs composants.

Les liens définissent une relation symétrique entre les objets. Ils ne sont pas pris en compte dans les algorithmes de classification et de filtrage. Par contre, toutes les autres fonctions les prennent en compte (contrôle du typage, migration des objets...).

Chaque point de vue représente une hiérarchie de classes définissant une taxonomie. Pour chaque classe d'un point de vue, il est possible de préciser ou de restreindre les attributs (clefs ou autres), mais il n'est pas possible d'en créer d'autres. La restriction peut être de plusieurs types :

- Définir un domaine (le co-domaine de l'attribut) ou ensemble des valeurs possibles,
- Définir des valeurs interdites,
- Définir une valeur par défaut,
- Définir une méthode d'assignation,
- Ajouter une fonction.

La notion de classe utilisée dans TROEPS est différente de la notion usuellement utilisée en programmation par objet. Aussi, la distinction entre concepts et classes dans TROEPS permet de séparer les aspects ontologiques (création et propriétés des objets) de leurs aspects taxonomiques (condition d'appartenance à une classe). Un exemple de cette séparation est donnée figure 1.4 pour le concept de MAISON et ses différents points de vue. La séparation classe/concept réduit l'importance des classes et permet d'envisager qu'un objet soit attaché à des classes différentes simultanément. Ainsi, chaque concept peut être vu sous des points de vue différents qui déterminent des taxonomies de classes différentes.

Enfin, il est possible de définir des passerelles entre certaines classes dans différents points de vue qui permettent alors d'inférer l'appartenance à une classe d'un point de vue donné sachant l'appartenance à une autre classe d'un autre point de vue. Dans l'exemple présenté figure 1.4, une passerelle peut être définie entre les classes PALACE du point de vue `Taille` et ELEVÉ du point de vue `Prix`. Ainsi d'un objet connu comme étant un

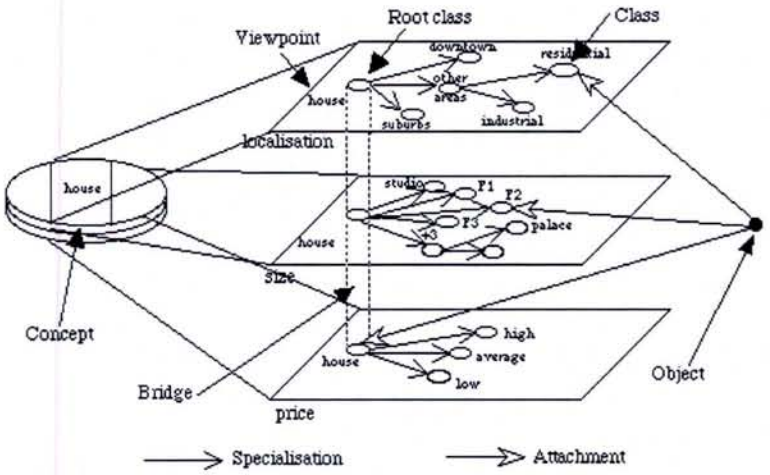


FIG. 1.4 – Le concept de MAISON (HOUSE) d'une base de connaissance TROEPS. Figure tirée de [48].

palace, on inférera que son prix est élevé.

1.3 Les logiques de descriptions

Les *logiques de descriptions* (LD) ou *logiques terminologiques* forment une famille de langages de représentation de connaissances inspirée de la logique des prédicats, des *réseaux sémantiques* et des *langages de schémas*. Elles trouvent leurs origines dans le développement du système KL-ONE [9]. La recherche sur les logiques de descriptions est très active et trouve un intérêt grandissant. Elles ont été utilisées dans de nombreuses applications telles que la modélisation de données conceptuelles [15], la manipulation de documents XML [14], la gestion des données [7], la gestion de requêtes [52], le traitement du langage naturel [104], les systèmes d'intégration d'informations [40].

Nous nous contenterons ici d'une présentation rapide qui pourra être bénéfiquement complétée par la lecture de l'ouvrage de référence [103], et des articles de synthèse comme [8, 101]. Nous présenterons plus avant le logiciel RACER que nous avons choisi d'intégrer dans notre système ROSA [89].

Les logiques de descriptions sont largement utilisées pour implanter des systèmes à bases de connaissances. Elles fournissent un modèle symbolique pour la représentation des connaissances du domaine d'une application. Ces logiques s'appuient sur les notions de concepts, de rôles et d'individus pour représenter les connaissances du domaine [101]. Un concept est un ensemble d'individus, alors qu'un rôle est une relation binaire entre individus. Toutes les logiques de descriptions possèdent les propriétés suivantes :

- Les concepts et les rôles sont définis à l'aide de constructeurs. Un concept se présente comme une conjonction et/ou une disjonction d'autres concepts, ou est défini à l'aide de rôles. Chaque rôle peut se voir associer des restrictions à l'instar des facettes associées à un attribut dans les langages de frames. Les restrictions portent généralement sur le co-domaine attaché à ce rôle (concept avec lequel le rôle établit

une relation), et la cardinalité du rôle, qui fixe le nombre minimal et/ou maximal de valeurs élémentaires i.e. d'individus que peut prendre le rôle. Une sémantique est associée à chaque description de concept et de rôle par l'intermédiaire d'une interprétation. Les manipulations opérées sur les concepts et les rôles sont réalisées en accord avec cette sémantique.

- Il existe deux niveaux de représentation des connaissances, le niveau terminologique (ou intensionnel) et le niveau factuel (ou extensionnel). Les concepts et les rôles sont définis et manipulés au niveau terminologique dans la TBox (*Terminology Box*). La description et la manipulation des individus s'effectue au niveau factuel dans la ABox (*Assertion Box*).
- Les concepts, et parfois les rôles, sont organisés hiérarchiquement selon la relation de subsomption : un concept C subsume un concept D si C est plus général que D au sens où l'ensemble des individus représentés par C contient l'ensemble des individus représentés par D.
- Il existe trois types de concepts : les concepts atomiques, les concepts primitifs et les concepts définis. Les concepts atomiques sont énoncés *ex nihilo* et servent à la construction des concepts primitifs et définis. Les propriétés d'un concept défini sont nécessaires et suffisantes pour qu'un individu qui les possède soit classé sous ce concept. En revanche, les propriétés d'un concept primitif sont nécessaires mais non suffisantes. Il n'est pas possible d'affirmer qu'un individu est un représentant d'un concept primitif en examinant ses seuls rôles.
- Les opérations à la base du raisonnement terminologique sont la *classification* et l'*instanciation*. La classification s'applique aux concepts et, dans certaines logiques de descriptions, aux rôles, et permet de déterminer la position d'un concept et d'un rôle dans leur hiérarchie respective. La construction et l'évolution de ces hiérarchies sont ainsi assistées par le processus de classification. L'instanciation permet de retrouver les concepts dont un individu est susceptible d'être une instance (classification d'instance). Cette notion d'instanciation est différente de celle utilisée pour les langages à objet, où l'instanciation correspond à la création d'un objet informatique à partir d'une classe.

Ainsi les logiques de descriptions peuvent être utilisées pour exprimer :

- Des connaissances *intensionnelles* sur des concepts et leurs relations,
- Des connaissances *extensionnelles* sous forme d'assertions sur les individus,
- Des *requêtes* sur les concepts et sur les individus.

1.3.1 Base terminologique

1.3.1.1 Syntaxe des concepts et des rôles

Les LD forment un sous-ensemble de la logique du premier ordre. Elles n'utilisent ni fonctions, ni variables explicites, mais des prédicats. Les prédicats unaires correspondent aux concepts, les prédicats binaires aux rôles. La description des concepts et des rôles est fondée sur l'utilisation de constructeurs. Le choix de ces constructeurs détermine le niveau d'expressivité et de performance de la procédure de classification de la famille de LD ainsi définie. Les premiers travaux théoriques ont porté sur une famille notée $\mathcal{FL}$ (*First*

Language). Le tableau 1.1 présente les différents constructeurs étudiés et les familles de langages de description qui en découlent.

constructeur	*syntaxe*	*langage*
concept	A, C	$\mathcal{FL}$
rôle	R	
conjonction de concepts	C $\sqcap$ D	
quantification universelle typée	$\forall R \cdot$ C	
quantification existentielle	$\exists R$	
top	$\top$	$\mathcal{AL}$ *
bottom	$\bot$	
négation ($\mathcal{C}$)	$\neg$A $\neg$C	
disjonction ($\mathcal{U}$)	C $\sqcup$ D	
quantification existentielle typée ($\mathcal{E}$)	$\exists R \cdot$ C	
conjonction de rôles ($\mathcal{R}$)	$R \sqcap S$	
cardinalité ($\mathcal{N}$)	($\geq$ nR) ($\leq$ nR) (= nR)	
hiérarchie de rôles	$R \sqsubseteq S$	$\mathcal{H}$
rôle inverse	R^{-1}	$\mathcal{I}$
cardinalité typée	($\geq$ n$R \cdot$ C) ($\leq$ n$R \cdot$ C) (= n$R \cdot$ C)	$\mathcal{Q}$

TAB. 1.1 – Tableau des constructeurs des logiques de descriptions et des familles de langages qui en découlent.

Le symbole A désigne un nom de concept atomique. R et S désignent des noms de rôle atomique. C et D sont des noms de concepts.

Le langage $\mathcal{FL}$ =$\{\text{A}, R, \text{C} \sqcap \text{D}, \forall R \cdot \text{C}, \exists R\}$ peut être enrichi par d'autres constructeurs pour former les langages suivants :

1. $\mathcal{AL} = \mathcal{FL} \cup \{\top, \bot, \neg\text{A}\}$
2. $\mathcal{ALC} = \mathcal{AL} \cup \{\neg\text{C}\}$
3. $\mathcal{ALN} = \mathcal{AL} \cup \{\geq \text{n}R, \leq \text{n}R\}$
4. $\mathcal{ALU} = \mathcal{AL} \cup \{\sqcup\}$
5. $\mathcal{ALE} = \mathcal{AL} \cup \{\exists R \cdot \text{C}\}$
6. $\mathcal{ALR} = \mathcal{AL} \cup \{R \sqcap \text{S}\}$
7. $\mathcal{ALC}_{R^+} = \mathcal{ALC} \cup \{R : R$ est un rôle atomique transitif$\}$[3]
8. $\mathcal{ALCN} = \mathcal{ALC} \cup \{\geq \text{n}R, \leq \text{n}R\}$
9. $\mathcal{ALCNR} = \mathcal{ALCN} \cup \{R \sqcap S\}$
10. $\mathcal{ALCH}_{R^+} = \mathcal{ALC}_{R^+} \cup \{R \sqsubseteq S\}$[4]
11. $\mathcal{ALCQHI}_{R^+}$ (ou $\mathcal{SHIQ}$) $= \mathcal{ALCH}_{R^+} \cup \{R^{-1}\}$[5] avec la cardinalité typée dénotée par $\mathcal{Q}$. Cette logique très expressive a été introduite dans [45].

[3] R est un rôle atomique transitif s'il n'est défini que par son nom et s'il est transitif.
[4] $R \sqsubseteq S$ veut dire que le rôle S subsume le rôle R : *objet* $\sqsubseteq$ *arc*.
[5] R^{-1} est le rôle inverse de R : *sommet-de*$^{-1}$ =*possede-sommet*.

Nous nous appuyons sur un extrait de la base de connaissances terminologiques (TBox) utilisée dans le cadre de nos travaux pour expliquer l'utilisation des principaux constructeurs des systèmes de logiques de descriptions disponibles. Les concepts et les rôles primitifs sont introduits par le symbole $\sqsubseteq$. Les concepts définis le sont par le symbole $\doteq$. La déclaration $C \sqsubseteq D$ construit le concept C comme étant subsumé par la description de D. La déclaration $C \doteq D$ définit le concept C comme étant équivalent à la description de D. Le tableau 1.2 peut être lu comme suit :

- Le concept TOP ($\top$) est un concept qui est à la racine de la hiérarchie des concepts. Il subsume tous les autres concepts. Tous les individus dans la base de connaissances sont des instances du concept TOP.
- Le concept BOTTOM ($\bot$) est subsumé par tous les concepts de la hiérarchie et il n'a pas d'instance.
- La négation de concept est définie par le constructeur $\neg$. Par exemple, ($\neg$RELATION) est un concept dont les instances sont tous les individus qui ne sont pas des instances du concept RELATION.
- La conjonction de concepts ($\sqcap$) définit l'intersection de concepts. L'expression (SOMMET $\sqcap$ $\neg$RELATION) définit le concept RÉGION comme la conjonction du concept SOMMET et de la négation du concept RELATION.
- La disjonction de concepts ($\sqcup$) représente l'ensemble des individus qui appartiennent au moins à l'un des concepts de la disjonction.
- La quantification existentielle typée : l'expression $\exists R \cdot C$ définit une contrainte sur les valeurs du rôle R ; le co-domaine est restreint au concept C, mais elle exige aussi l'existence d'au moins un couple d'individus `(x,y)` liés par l'intermédiaire de R, tel que `y` est une instance de C. Par exemple, le concept RELATION définit un ensemble d'individus qui ont (au moins) un *arc* qui est de type RÉGION. La quantification existentielle simple est l'équivalent d'une quantification existentielle typée avec le concept TOP comme restriction du co-domaine du rôle R.
- La quantification universelle typée : une contrainte sur le co-domaine des valeurs d'un rôle est introduite par l'expression $\forall R \cdot C$, le co-domaine du rôle R est restreint au concept C. Contrairement à l'expression $\exists R \cdot C$, cette expression impose une contrainte sur le co-domaine de R mais elle n'exige pas l'existence d'individus en relation par le rôle. La quantification universelle simple est l'équivalent d'une quantification universelle typée avec le concept TOP comme restriction du co-domaine du rôle R.
- La cardinalité typée : les constructeurs $\geq \mathtt{n} R \cdot C$, $\leq nR \cdot C$ et $= \mathtt{n} R \cdot C$ fixent à `n` le nombre minimal, maximal ou exact des individus `y` auxquels un individu `x` est lié par le rôle R en restreignant le co-domaine du rôle à un concept particulier C. Une cardinalité simple est équivalente à une cardinalité typée avec le concept TOP comme restriction du co-domaine du rôle R. Par exemple, l'expression ($= 3arc \cdot$ RÉGION) définit des relations ayant trois *arc* reliant trois RÉGION, i.e. des relations ternaires.

1.3.1.2 Sémantique des concepts et des rôles

La sémantique associée aux logiques de descriptions est définie par une *interprétation* $(\Delta^I, .^I)$. Δ^I est un ensemble non vide qui désigne un domaine d'interprétation et $.^I$ est une

la base des concepts
$\textsc{Sommet} \sqsubseteq \textsc{Top}$
$\textsc{Relation} \sqsubseteq \textsc{Sommet} \sqcap (\exists arc \cdot \textsc{Région})$
$\textsc{Région} \sqsubseteq \textsc{Sommet} \sqcap \neg\textsc{Relation}$
$\textsc{Région} \sqcup \textsc{Relation}$
$\textsc{RelationBinaire} \doteq \textsc{Relation} \sqcap (= 2arc \cdot \textsc{Région})$
$\textsc{RelationTernaire} \doteq \textsc{Relation} \sqcap (= 3arc \cdot \textsc{Région})$
$arc \sqsubseteq toprole$
$objet \sqsubseteq arc$
$sujet \sqsubseteq arc$

TAB. 1.2 – Extrait de la base des concepts utilisés dans le cadre de ce travail.

fonction d'interprétation. Par l'intermédiaire de la fonction $.^I$ les concepts sont interprétés comme des sous-ensembles de Δ^I, et les rôles sont interprétés comme des sous-ensembles du produit $\Delta^I \times \Delta^I$. Un *modèle* d'un concept C est une interprétation où C^I (l(image de C par I) est un ensemble non vide. Un concept C est *satisfiable* s'il a un modèle.

La sémantique associée aux constructeurs définis dans le tableau 1.1 est donnée dans le tableau 1.3.

syntaxe	*sémantique*
$\mathtt{A}$	$\mathtt{A}^I \subseteq \Delta^I$
R	$R^I \subseteq \Delta^I \times \Delta^I$
$\mathtt{C} \sqcap \mathtt{D}$	$\mathtt{C}^I \cap \mathtt{D}^I$
$\forall R \cdot \mathtt{C}$	$\{X \in \Delta^I \mid \forall Y.(X,Y) \in R^I \Rightarrow Y \in \mathtt{C}^I\}$
$\exists R$	$\{X \in \Delta^I \mid \exists Y.(X,Y) \in R^I\}$
$\top$	Δ^I
$\bot$	ϕ
$\neg\mathtt{C}$	$\Delta^I \backslash \mathtt{C}^I$
$\mathtt{C} \sqcup \mathtt{D}$	$\mathtt{C}^I \cup \mathtt{D}^I$
$\exists R \cdot \mathtt{C}$	$\{X \in \Delta^I \mid \exists Y.(X,Y) \in R^I \wedge Y \in \mathtt{C}^I\}$
$R \sqsubseteq S$	$R^I \cap S^I$
$(\geq \mathtt{n}R)(\leq \mathtt{n}R)(= \mathtt{n}R)$	$\{X \in \Delta^I \mid \#\{Y \mid (X,Y) \in R^I\} \geq, (\leq, =)\mathtt{n}\}$
R^{-1}	$\{(Y,X) \mid (X,Y) \in R^I\}$
$(\geq \mathtt{n}R \cdot \mathtt{C})(\leq \mathtt{n}R \cdot \mathtt{C})(= \mathtt{n}R \cdot \mathtt{C})$	$\{X \in \Delta^I \mid \#\{Y \mid (X,Y) \in R^I \wedge Y \in \mathtt{C}^I\} \geq, (\leq, =)\mathtt{n}\}$

TAB. 1.3 – Sémantique de constructeurs.

D'une façon intuitive, les interprétations de constructeurs sont à comprendre comme suit. L'interprétation d'un concept primitif A est un sous-ensemble du domaine d'interprétation Δ^I. Pour un rôle R, l'interprétation est un sous-ensemble de $\Delta^I \times \Delta^I$. L'interprétation de la conjonction (respectivement de la disjonction) de concepts est l'intersection (respectivement l'union) de leurs interprétations. La quantification universelle typée est interprétée par un sous-ensemble de Δ^I dont les individus $\mathtt{x}$ sont liés par le rôle R à des individus $\mathtt{y}$ tel que $\mathtt{y} \in \mathtt{C}^I$. Par contre, la quantification existentielle typée est interprétée

par un sous-ensemble de Δ^I dont les éléments x ont (au moins) un lien par l'intermédiaire du rôle R à d'autres éléments y tel que $\mathtt{y} \in \mathrm{C}^I$. De la même façon, l'interprétation de la quantification existentielle est un sous-ensemble de Δ^I dont les éléments sont liés par R à (au moins) un élément dont le type n'est pas précisé (une instance du concept $\top$). L'interprétation de la négation de C est l'ensemble complémentaire de C^I.

1.3.1.3 Raisonnement dans la base des concepts

Dans la partie intensionnelle de la base de connaissances (TBox), trois types de raisonnements sont effectués : le test de *subsomption*, le test de *satisfiabilité*, et la *classification*. Le test de subsomption est un mécanisme sur lequel peuvent se baser les deux autres mécanismes.

1. Déterminer si le concept D subsume le concept C revient à vérifier que le concept D (le *subsumant*) est plus général que C (le *subsumé*). Autrement dit, la subsomption revient à vérifier que l'ensemble des instances de C est un sous-ensemble de celui de D. Formellement, $\mathrm{C} \sqsubseteq \mathrm{D}$ si et seulement si $\mathrm{C}^I \subseteq \mathrm{D}^I$ pour toute interprétation I.
2. Vérifier qu'un concept D est *satisfiable* revient à vérifier qu'il a une interprétation I telle que $\mathrm{D}^I \neq \phi$; sinon D est non satisfiable. La non satisfiabilité est un cas particulier de subsomption où le subsumant est le concept $\bot$. Si $\mathrm{D} \sqsubseteq \bot$, alors D est non satisfiable.
3. La classification est le processus de base pour construire la hiérarchie des concepts. La classification consiste à positionner un concept par rapport aux autres concepts de la hiérarchie. La relation de subsomption entre le concept à positionner et tous les concepts dans la hiérarchie est vérifiée, puis le concept est placé entre ses subsumants les plus spécifiques et ses subsumés les plus généraux. Le concept D est un subsumant plus spécifique du concept C si et seulement si $\mathrm{C} \sqsubseteq \mathrm{D}$ et s'il n'existe pas un autre concept F tel que $\mathrm{C} \sqsubseteq \mathrm{F} \sqsubseteq \mathrm{D}$. Pour les subsumés, D est un subsumé plus général de C si et seulement si $\mathrm{D} \sqsubseteq \mathrm{C}$ et s'il n'existe pas un autre concept F tel que $\mathrm{D} \sqsubseteq \mathrm{F} \sqsubseteq \mathrm{C}$.

1.3.2 Base de faits

1.3.2.1 Assertions

La base des assertions contient les connaissances extensionnelles concernant le domaine de l'application. Les assertions concernant l'appartenance des individus aux concepts sont appelées *instanciations de concepts*, elles sont de la forme C(a), où l'individu a appartient au concept C. De la même façon, les assertions concernant les relations entre les individus sont des *instanciations de rôles*. Elles sont de la forme $R(\mathtt{a}, \mathtt{b})$ où les individus a et b sont liés par le rôle R.

Le tableau 1.4 présente des faits extraits d'une base de connaissances construite au cours de notre travail. Les faits présentés définissent des RÉGIONS dénommées champ, pré et bois ainsi qu'une RELATION nommée isole. La relation isole est précisée par l'instanciation des rôles *objet* et *sujet* définis dans le tableau 1.2. Cette description représente la phrase "Le pré isole le champ du bois".

la base d'assertions
RÉGION (champ)
RÉGION (pré)
RÉGION (bois)
RELATION (isole)
objet (champ, isole)
objet (bois, isole)
sujet (pré, isole)

TAB. 1.4 – Tableau des assertions permettant la définition des individus de type RÉGION : champ, pré et bois et d'un individu de type RELATION : isole. Isole est défini comme ayant pour *sujet* l'individu pré et comme *objet* les individus champ et bois.

1.3.2.2 Raisonnement dans la base des assertions

D'une façon similaire à la base des concepts, plusieurs types de raisonnements, qui s'appuient sur l'*instanciation*, sont possibles dans la base des assertions. Nous détaillons à présent ces types de raisonnements :

1. L'*instanciation* consiste à vérifier qu'un individu x est une instance du concept C.
2. La *réalisation* permet de trouver les concepts les plus spécifiques dont un individu x est une instance.
3. Le *calcul de l'extension* consiste à trouver toutes les instances d'un concept C.

Les types de raisonnements dans la base des concepts et la base des assertions ne sont pas indépendants les uns des autres. Par exemple :

- Instanciation et satisfiabilité : l'individu x est une instance du concept C dans une base de connaissances Σ ($\Sigma \models C(x)$) si et seulement si Σ étendue par $\neg C(x)$ ($\Sigma \cup \neg C(x)$) n'est pas satisfiable.
- Cohérence et satisfiabilité : le test de la satisfiabilité de la base de connaissances se ramène à vérifier que cette base est cohérente (Σ est satisfiable $\Longleftrightarrow \forall x : \Sigma \cup \bot(x)$ n'est pas satisfiable).

La sémantique des assertions d'instanciation de concepts et des assertions d'instanciation de rôles est donnée comme pour la base de concepts par un domaine (Δ^I) et une fonction ($.^I$) d'interprétation. La fonction $.^I$ associe, sous l'hypothèse du nom unique, une interprétation a^I de Δ^I à chaque individu a. L'hypothèse du nom unique assure que si a et b sont des noms d'individus différents, alors $a^I \neq b^I$.

Un modèle d'une base de connaissances Σ est une interprétation $I = (\delta, .^I)$ où chaque déclaration dans la base des concepts et la base des assertions est satisfiable.

1.3.3 Systèmes de logiques de descriptions

Le rapport entre l'expressivité des langages des logiques de descriptions et la complexité des raisonnements a mené au développement de nombreux systèmes de logiques de descriptions [47] qui peuvent être regroupés en trois catégories. Dans ces systèmes, les services de raisonnement fournis à l'utilisateur sont largement influencés par l'ensemble

des constructeurs choisis pour construire les langages [85]. Dans la première catégorie, l'ensemble de constructeurs est restreint de façon à ce que la relation de *subsomption* soit calculée efficacement et si possible en temps polynomial, ce qui constitue une approche limitative mais complète ; le système CLASSIC [8] en est l'exemple le plus significatif. La deuxième catégorie peut être décrite comme expressive mais incomplète. L'incomplétude des algorithmes de raisonnement est due à l'expressivité du langage. Les systèmes LOOM [78] et BACK [105] en sont des exemples. Après avoir étudié les causes de l'incomplétude qui peuvent être rapportées aux combinaisons de constructeurs, des algorithmes complets ont été réalisés. Les systèmes de ce type sont décrits comme expressifs et complets. KRIS [2], RACER [41, 42] et FACT [45] suivent cette dernière approche. Le système LOOM est encore utilisé dans certaines applications, RACER et FACT peuvent être considérés comme les deux systèmes les plus opérationnels actuellement.

1.3.4 Le système Racer

RACER est un système de LD qui implémente le langage $\mathcal{ALCQHI}$ (appelé $\mathcal{SHIQ}$). C'est le langage $\mathcal{ALC}$ étendu par les contraintes de cardinalité, la hiérarchie de rôles, les rôles inverses et les rôles transitifs. RACER est le premier système de LD qui implémente $\mathcal{SHIQ}$ avec une base de concepts (la TBox) et une base de faits (la ABox). Il a été développé à l'université de Hamburg en Allemagne[6]. RACER peut gérer plusieurs bases de concepts et plusieurs bases de faits indépendantes les unes des autres en même temps. De plus, il est doté d'une "interface de programmation d'application" (API[7]) qui permet d'accéder, depuis une application Java, d'une façon interactive, aux différentes fonctions contenues dans sa librairie.

L'ensemble des constructeurs implémentés par RACER est présenté dans la table 1.5. Les définitions des concepts en RACER peuvent être cycliques : elles peuvent se référencer directement (le concept A fait référence au concept B et B fait référence à A) ou indirectement (ex : A fait référence à B, B fait référence à C et C fait référence à A). RACER peut gérer des rôles transitifs et les rôles peuvent être organisés dans une hiérarchie par une relation de subsomption. De plus, la dernière version de RACER prend en compte les domaines concrets qui définissent des types de base comme *entier*, *réel* ou *chaîne de caractères*.

1.4 La représentation des relations

Les relations apparaissent dans tous les systèmes de représentation de connaissances. Cependant, cette notion n'est pas toujours bien maîtrisée. En fait, différents types de relations existent qui sont plus ou moins bien étudiées et/ou formalisées. Nous nous intéressons aux relations qui interviennent dans les structures spatiales. Nous allons approfondir cette notion afin de comprendre ce qu'est une relation et quelles connaissances portant sur les relations peuvent être représentées. En particulier, nous souhaitons pouvoir les décrire et les classer.

[6] `http://www.sts.tu-hamburg.de/~r.f.moeller/racer/`
[7] Application Programming Interface

Notation LD	*Syntaxe* RACER
$\neg$ C	(not C)
$C_1 \sqcap ... \sqcap C_n$	(and C_1...C_n)
$C_1 \sqcup ... \sqcup C_n$	(or C_1...C_n)
$\exists R \cdot C$	(some R C)
$\forall R \cdot C$	(all R C)
$\leq \mathtt{n} R$	(at-most n R)
$\geq \mathtt{n} R$	(at-least n R)
$= \mathtt{n} R$	(exactly n R)
$\leq \mathtt{n} R \cdot C$	(at-most n R C)
$\geq \mathtt{n} R \cdot C$	(at-least n R C)
$= \mathtt{n} R \cdot C$	(exactly n R C)
$C \sqsubseteq D$	(implies C D)
$C \doteq D$	(equivalent C D)

TAB. 1.5 – Syntaxe des constructeurs du système RACER.

1.4.1 Différents types de relations

1.4.1.1 Spécialisation

La relation *est-un* (ou *sorte-de*) est une relation d'ordre partiel : transitive, anti-symétrique et réflexive. Elle est à la base des systèmes de représentation de connaissances. C'est à la fois la relation la mieux connue et la mieux formalisée. Cette relation permet de définir des hiérarchies de spécialisation pour construire des taxonomies.

1.4.1.2 Partie-de / Partie-Tout

La relation *partie-de* est le deuxième type de relation étudiée. Cette relation peut avoir plusieurs sens [125] :

- La notion d'objets composés de plusieurs autres objets : une voiture est composée de roues, d'une carrosserie, etc,
- La notion de membre – collection : une vache et son troupeau,
- La notion de portion – masse : une part de tarte,
- La notion de matière – objet : l'acier entre dans la construction d'une voiture,
- La notion spatiale : une oasis fait partie du désert,
- La notion de structure – activité : le fait de payer fait partie de l'action de faire les courses.

Cette relation peut définir une hiérarchie sur les objets. Elle a été implantée, par exemple, dans YAFOOL [31] sous le nom *composé-de*. Le sens de la relation *composé-de* se réfère à la notion d'objets composés mais elle peut être aisément utilisée pour une relation de membre-collection. C'est aussi celle considérée dans TROEPS [48].

1.4.1.3 Relations spatiales

Dans le cadre de notre travail, nous nous intéressons plus particulièrement aux relations spatiales pour représenter l'organisation spatiale des exploitations agricoles que nous étudions. On distingue habituellement trois types de relations spatiales :

- Topologie : notion de partie-tout, de connexion. Les modèles sont la topologie mathématique, la méréologie[8][125] et plus récemment les théories algébriques comme les théories RCC-8 ou RCC-5 [112].
- Distance : notion de proximité – éloignement. La distance entre deux objets est évaluée dans un cadre de référence. Par exemple, un champ donné peut être loin des bâtiments d'exploitation pour le déplacement des animaux alors qu'il pourra être considéré comme proche vis-à-vis d'un déplacement en tracteur. De même pour deux champs à égale distance métrique, celui dont l'accès nécessite la traversée d'une agglomération sera considéré comme étant plus loin que celui accessible facilement.
- Orientation : positionnement tel que à gauche / à droite, devant / derrière ou orientation cardinale. Ici aussi intervient un cadre ou un objet de référence. Par exemple pour l'expression : "le pré est derrière la ferme", le cadre est donné par l'orientation intrinsèque du bâtiment (il existe un avant et un arrière).

Les plus étudiées de ces relations sont les relations topologiques. En effet, ce sont des relations binaires bien identifiées ne faisant pas appel à un cadre ou objet de référence. Une relation d'orientation ou de distance est souvent ternaire pour prendre en compte la référence qui donne tout son sens à la relation [44]. Les relations topologiques ont fait l'objet de nombreux travaux de formalisation, tandis que les formalismes des relations de distances et d'orientation restent incomplets et sont donc moins facilement utilisables [123]. En particulier, la théorie RCC-8 [112] définit les relations topologiques sur la notion de connexion entre régions pour des objets connexes et fermés. Nous ne détaillons pas plus ces formalismes ici, mais il est possible de se référer à [121] pour une présentation générale des travaux sur les formalisations de l'espace.

1.4.2 Représentation des relations

Dans les réseaux sémantiques, les concepts sont décrits par les relations qu'ils possèdent avec les autres concepts. La relation est donc fondamentale mais elle est binaire et elle est un simple lien entre concepts.

Dans les langages de *frames*, une relation permet de créer une dépendance entre deux frames. La relation et sa relation inverse sont définies grâce à deux liens orientés, réciproques l'un de l'autre. Chaque lien est représenté par un attribut muni d'une facette de typage et d'une facette lien-inverse. Le frame de départ correspond au domaine de la relation, tandis que le frame d'arrivée, dans lequel la relation prend ses valeurs, s'appelle le co-domaine de la relation [100].

Dans les graphes conceptuels [120, 20], la relation tient une place importante dans la définition des connaissances. Les relations sont des structures typées reliant des concepts et elles sont regroupées en classes. À l'origine, les relations ne forment pas une hiérarchie, mais il existe une extension du formalisme initial des graphes conceptuels permettant

[8]Une théorie de la relation "partie – tout".

de spécialiser des relations. Les relations sont n-aires et sont représentées par des profils permettant de désigner le type de chaque concept relié par la relation [70].

Les premiers langages à objets offrent la manipulation de relations binaires [83]. Ce sont en fait des liens orientés ou des pointeurs d'une classe vers une autre. La relation n'est pas réifiée i.e. elle n'est pas représentée par une classe. Elle possède un nom par lequel on peut accéder à la valeur, et il est possible de typer la valeur de la relation.

Par la suite, des relations n-aires ont été définies [114]. Une telle relation est alors définie par une liste ordonnée de classes d'objets : le co-domaine. Le nombre de classes constitue le *degré* de la relation. Chaque classe est représentée par un champ identifié par sa position et auquel on peut assigner un unique nom de *rôle*. La valeur de la relation est un ensemble de tuples qui instancient les classes de la relation. On peut associer à chaque champ une cardinalité. Celle-ci restreint le nombre de tuples sur une instance pour ce champ. Les opérations sur les relations sont l'ajout, la suppression, la modification d'un tuple et des requêtes telles que le test d'appartenance d'un tuple, ou les tuples présents pour un champ donné d'une instance.

Pour les logiques de descriptions, une relation binaire correspond à l'instanciation d'un rôle comme nous l'avons présenté précédemment. Une relation associe deux objets qui existent effectivement ou bien fixe une cardinalité pour un rôle donné. Un rôle représente une propriété associée à un concept. Les relations n-aires ont été étudiées dans [25, 13]. Mais le résultat de ces travaux théoriques n'est pas encore disponible dans des systèmes opérationnels comme RACER ou FACT.

1.4.3 Les relations vues comme des concepts

Pour représenter les relations, on peut faire appel au modèle entité-association manipulé dans les systémes de bases de données ou encore dans les modèles OMT ou UML. Mais dans les bases de données relationnelles, tout est relation et c'est la logique des prédicats qui est utilisée pour la manipulation des données. Dans les systèmes de RCO en revanche, tout est objet. Peu de logiciels mettent en œuvre à la fois les entités et les relations (un exemple : ILOG power classes). Les relations, dans les systèmes de RCO, sont en général représentées soit sous forme d'attribut-lien (attribut ayant pour type la classe cible de la relation), soit sous forme de classe.

Les relations sont une spécialité du système de RCO Y3 [31]. Y3 est un système à objets qui se compose essentiellement du langage hybride YAFOOL et de l'interface YAFEN. En Y3, tout objet, classe ou instance, est un *frame*, constitué d'un ensemble de *slots* représentant aussi bien des attributs que des méthodes. Tout attribut peut être muni de facettes déclaratives ou procédurales. Attribut, facettes et méthodes sont des objets à part entière. En particulier, les relations (qui sont des attributs particuliers) sont réifiées et héritent d'une classe RELATION elle-même spécialisation de la classe ATTRIBUT.

Du fait des spécificités de YAFOOL, il a été possible d'étendre la représentation des relations et de s'approcher de la représentation des logiques de descriptions par adjonction de facettes pour obtenir ainsi un langage d'une grande expressivité, mais d'une certaine complexité [66, 80].

Mais un tel système reste adapté aux seules relations binaires. Aussi, d'autres recherches dans le cadre des systèmes de RCO ont mené au développement du système

AROM : *Allier Relation et Objets pour Modéliser* [109]. Dans ce système, les liens entre objets sont explicitement représentés par des entités nommées associations. De plus, ce système propose une représentation graphique des connaissances, proche des notations UML qui a l'avantage de faciliter la communication parmi les concepteurs et utilisateurs de systèmes à bases de connaissances. Il offre aussi un *langage de modélisation algébrique* (LMA) pour l'expression des connaissances structurelles (équations et contraintes).

AROM est constitué de deux niveaux. Un premier, le noyau, appelé AROM*min* permet la définition de bases de connaissances constituées des classes avec leurs instances, d'une part, et des relations avec leurs tuples (instances d'association ou n-uplet), d'autre part. Les classes comme les relations sont organisées en hiérarchies de spécialisation, mais ces spécialisations n'obéissent pas aux mêmes règles. Le deuxième niveau intègre le langage de modélisation algébrique permettant de faire de la simulation ou de la satisfaction de contraintes en relation avec la base de connaissances gérée par le noyau.

Dans AROM, les associations jouent un rôle aussi important dans la représentation de connaissances que les classes. Elles sont semblables aux associations du modèle UML : une association représente un ensemble de liens entre n ($n \geqslant 2$) classes, distinctes ou non. Un lien est un n-uplet d'objets appartenant aux extensions (ensembles d'instances) des classes reliées par l'association. Chaque association possède un nom. Une association est décrite par ses rôles, ses variables et ses contraintes. Un rôle r d'une association correspond à une connexion entre l'association et une des classes connectées, appelée classe correspondante du rôle. Chaque association n-aire possède donc n rôles et la valeur de chacun est une instance de la classe correspondante. Chaque rôle possède un nom et une multiplicité. La multiplicité d'un rôle a le même sens qu'en UML. Il s'agit d'un intervalle d'entiers. La multiplicité est décrite par la facette *multiplicity* : pour laquelle est donnée la valeur minimum *(min :)* et la valeur maximum *(max :)*. Pour cette dernière, on utilise, comme en UML, le symbole $*$ pour dénoter une valeur infinie. Outre la facette *multiplicity* :, la facette documentation : permet d'associer une documentation à un rôle.

Les associations sont organisées en hiérarchies grâce à une relation de spécialisation. La spécialisation d'associations, comme celles des classes, est simple en AROM. Elle permet de greffer sur une hiérarchie d'associations un héritage de rôles, de variables et de facettes. Une association A2 spécialise une association A1 par les opérations suivantes :

- Spécialisation de la classe correspondante (co-domaine) d'un rôle de A1 ;
- Modification ou ajout d'une facette à une variable ou à un rôle de A1 ;
- Ajout d'une contrainte à une variable ou à un rôle de A1 ;
- Ajout d'une contrainte d'association à A1 ;
- Ajout d'une variable à A1.

La spécialisation d'associations correspond à l'inclusion ensembliste des liens. Autrement dit, les liens d'une association appartiennent à l'ensemble des liens de sa super-association, si elle existe. Par conséquent, la spécialisation d'associations par ajout de rôle n'est pas autorisée en AROM : elle reviendrait à considérer l'inclusion de n+1-uplets dans un ensemble de n-uplets. L'arité d'une association est donc préservée dans toute sa hiérarchie.

1.5 Conclusion

Nous avons présenté dans ce chapitre, différents langages développés pour répondre aux besoins de représentation de connaisances. Nous avons particulièrement insisté sur les derniers travaux sur les logiques de descriptions et les systèmes de représentation de connaissances par objets qui sont développés actuellement.

Nous souhaitons représenter des structures spatiales, et pour ce faire nous avons particulièrement besoin d'exprimer des relations entre des entités. Des systèmes présentés, il apparaît deux types d'approches différentes pour la représentation des relations :

- Certains placent les relations au même niveau que les entités. C'est notamment le cas pour YAFOOL, de façon encore plus marquée avec AROM.
- D'autres présentent les relations comme des liens particuliers entre des concepts : les LD ou les graphes conceptuels.

De tous les systèmes présentés, ce sont les logiques de descriptions qui offrent le plus d'expressivité avec des résultats théoriques importants sur la complexité et la décidabilité en fonction des constructeurs du langage considéré. Des travaux portent sur la comparaison des capacités de représentation des différents formalismes. Les résultats obtenus réduisent les différences au point que certains se demandent s'il existe vraiment des différences.

Actuellement, le choix du formalisme de représentation pour notre système met en balance un système qui représente les relations comme les entités : AROM et un système très expressifs : le système de LD RACER. AROM est encore en phase de prototypage alors que RACER offre une des LD les plus expressives tout en gérant les deux niveaux de raisonnement terminologique et factuel. De plus, il est possible simplement de définir des relations sous forme de concepts en détournant légèrement la notion de rôle, ce qui réduit l'écart entre les deux systèmes. Notre choix pour le formalisme de représentation et du système de raisonnement associé s'est donc porté sur l'utilisation du système de LD RACER. Reste la difficulté de manipuler des graphes, qui est au cœur du travail présenté ici.

Chapitre 2

Raisonnement à Partir de cas

Sommaire

2.1 Le raisonnement par analogie 30
2.1.1 Principe 30
2.1.2 Notion de similarité 31
2.2 Le raisonnement à partir de cas 32
2.2.1 Le modèle 33
2.2.2 Le cycle du RàPC 34
2.3 Le RàPC en milieu hiérarchique 36
2.3.1 L'indexation 36
2.3.2 La remémoration 37
2.3.3 La classification dure, les chemins de similarité et l'adaptation 37
2.3.4 La classification élastique 39
2.3.5 Les difficultés de l'adaptation 41
2.4 Le raisonnement à partir d'explications 41
2.5 Conclusion 43

L'étude des raisonnements est une préoccupation qui a mobilisé les chercheurs bien avant l'arrivée des ordinateurs. On peut faire remonter cette recherche à Aristote, qui le premier formalisa un mode de raisonnement fondé sur la logique : la syllogistique. Ce sont d'ailleurs les raisonnements logiques qui furent les plus étudiés au cours du temps. Parmi les noms célèbres, nous pouvons citer Pascal, Liebniz, Boole, Frege. Ainsi à l'avènement de l'informatique, les raisonnements logiques déductifs furent rapidement utilisés et ont notamment servi de base au développement des systèmes à base de règles (ou systèmes experts) fondés sur des connaissances de la forme "si...alors".

Les systèmes fondés sur le raisonnement déductif ont l'intérêt scientifique de la rigueur. Mais ces systèmes ont vite montré leurs faiblesses dues à leur trop grande rigidité. En effet, si les modes de raisonnement humain s'accommodent du raisonnement déductif, notamment dans les disciplines mathématiques, ils ne s'y limitent pas. Aussi, l'étude de systèmes mettant en œuvre des raisonnements non rigoureux, comme les raisonnements inductif et/ou abductif n'est pas dénuée d'intérêt. Ces sont ces types de raisonnements qui sont utilisés dans les raisonnements par analogie, un des modes de raisonnement les plus utilisés par l'homme. Pour de nombreux théoriciens, l'analogie est une des composantes clés de l'intelligence. Les mécanismes d'analogie se retrouvent dans d'autres types de raisonnement et notamment dans le raisonnement par classification [43].

Le raisonnement par analogie est sûrement le moins connu, le moins bien compris. Mais les recherches tant en psychologie cognitive que dans le cadre des systèmes informatiques se développent. En informatique, un cas particulier de raisonnement par analogie, le raisonnement à partir de cas, est maintenant assez largement étudié [68]. C'est ce type de raisonnement que nous étudions dans le cadre du développement de notre système ROSA.

Dans ce chapitre, après avoir présenté le raisonnement analogique et la notion de similarité qui le sous-tend, nous exposons le raisonnement à partir de cas. Nous en détaillons le cycle puis nous présentons un modèle du raisonnement à partir de cas en milieu hiérarchique avant de conclure.

2.1 Le raisonnement par analogie

2.1.1 Principe

Le raisonnement analogique repose sur le transfert et l'adaptation d'une propriété ou d'une relation, notée β, existant entre les objets "d'un univers" (la source), sur les objets "d'un autre univers" (la cible). Il met en œuvre un mécanisme de mise en correspondance, ou projection (*mapping*) des objets de la source et de la cible afin de transposer des connaissances d'un univers d'application à un autre. Ce transfert se fait suivant un certain point de vue : un but à atteindre ou un problème à résoudre. La question essentielle est de déterminer dans quelle mesure la propriété analogique β (relation ou dépendance) à transférer et la relation de ressemblance α (ou correspondance) entre les deux univers permettent la transposition. La formalisation de ce processus permet de définir le *paradigme d'analogie* proposé par [21] et présenté figure 2.1.

Par exemple, on peut trouver une analogie entre un circuit électrique et un circuit

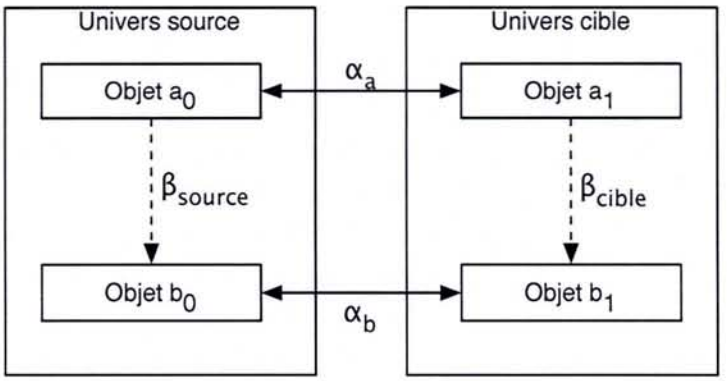


FIG. 2.1 – Le paradigme d'analogie. α représente la relation de ressemblance entre les univers sources (0) et les cibles (1). β représente la propriété analogique transférée.

hydraulique. L'univers source est constitué des objets : pile, résistance, fils, tension et courant. Ces objets sont mis respectivement en correspondance avec ceux de l'univers cible : pompe, filtre, tuyaux, pression et débit.

Le raisonnement par analogie ne prétend pas toujours trouver une solution, ni même, lorsqu'une solution est proposée, que c'est une bonne solution. Il suppose l'existence d'une ressemblance, d'une similitude entre les objets des différents univers. Cette similitude est mise à profit pour établir le transfert de propriétés d'un univers à l'autre. Dans l'exemple proposé ci-avant, on peut envisager de calculer la perte de charge dans le filtre sachant comment calculer la différence de potentiel aux bornes de la résistance. En effet, la similitude apparente entre les deux univers est grande. Mais la différence de nature des fluides rend plus délicate le transfert de la propriété et il faut des connaissances supplémentaires pour permettre une adaptation de la relation U=RI dans l'univers des circuits hydrauliques.

2.1.2 Notion de similarité

Le principal problème du raisonnement par analogie consiste à déterminer les ressemblances existantes entre le problème à résoudre (la cible) et des problèmes déjà résolus parmi lesquels on cherche un exemplaire qui pourrait aider à sa résolution.

Dans l'exemple proposé, il y a une similarité "évidente" (ou superficielle) sur les notions de circuit et de fluide. La mise en correspondance terme à terme (ou correspondance en profondeur) des objets dans les univers vient ensuite et confirme l'impression de départ. C'est d'ailleurs un des atouts de l'analogie que de faire apparaître comme évidente une similarité au premier coup d'œil. Les expériences menées par Eynard et Ripoll [35] semblent bien démontrer que l'esprit humain s'appuie bien sur des indications superficielles et non sur une mise en correspondance profonde pour établir l'analogie.

Pourtant, la similarité est souvent envisagée pratiquement comme une mise en correspondance des descripteurs de la cible avec ceux des problèmes sources. On définit alors la similarité comme étant une distance entre les problèmes. Plus la similarité est forte

plus la distance calculée est faible. Il s'agit de déterminer les "bons" descripteurs dont les valeurs sont identiques ou pour le moins très proches en ayant dans l'idée que la solution d'un problème est d'autant plus simple à trouver qu'il ressemble à celui déjà résolu.

Pour peu que la description des problèmes soit complexe, la mise en correspondance de leurs descripteurs peut être coûteuse en temps de calcul. Et cette conception de la similarité est alors en contradiction avec l'idée que l'analogie doit être considérée comme une stratégie permettant de réduire la combinatoire. Mais son avantage tient au fait qu'il est alors souvent possible de mesurer une distance pour chacun des descripteurs entre les problèmes source et cible. Ce calcul permet alors de classer les problèmes sources par rapport au problème cible à résoudre. Il peut aussi servir l'adaptation en considérant l'écart existant entre les descripteurs du problème source choisi et ceux du problème cible dont on cherche à calculer la solution.

2.2 Le raisonnement à partir de cas

Le *raisonnement à partir de cas* (RÀPC), ou fondé sur des cas (*case-based reasoning*) est un cas particulier de raisonnement par analogie où les univers source et cible sont confondus (figure 2.2). Plutôt que de transférer une propriété d'un univers dans un autre, le RÀPC cherche à adapter l'application de la propriété à transférer sur les objets d'une cible afin de construire des connaissances nouvelles sur celle-ci [94]. Il est fondé sur l'utilisation d'unités d'expériences appelées *cas*, constituées classiquement de problèmes résolus. Il consiste à trouver une solution à un problème, appelé "problème cible" en s'appuyant sur une base de cas constituée des unités d'expériences.

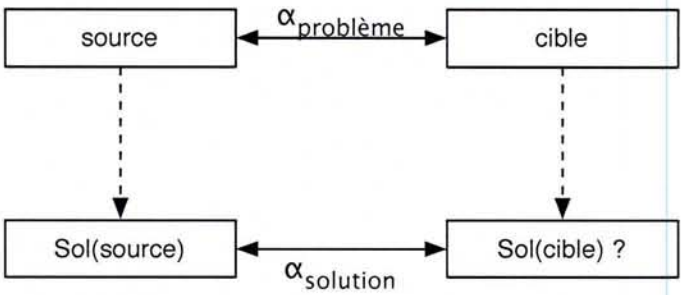


FIG. 2.2 – Ce schéma peut se lire : sachant que `Sol(cible)` est à `Sol(source)` ce que `cible` est à `source`, quelle valeur proposer pour `Sol(cible)` étant donné `source`, `Sol(source)` et `cible` ? [72]

Ce domaine de recherche a pour origine les travaux de Schank [116] qui cherchait à définir la notion de système intelligent. Dans les années 70, ce dernier travaillait sur la compréhension du langage naturel et sur le rôle de la mémoire dans le raisonnement humain. Il proposa alors le concept de script permettant de décrire des comportements stéréotypiques. Une fois reconnu dans une situation à analyser, ces scripts servent à expliquer, prédire la suite des actions qui doivent avoir lieu. Un script devenu célèbre est le

script des actions attendues pour manger dans un restaurant : entrer, s'asseoir, prendre et lire la carte, commander, être servi, manger, payer, sortir.

Au cours des années 80, Schank définit le script comme une description très personnelle d'une situation de la vie courante, fruit de l'apprentissage au cours de nos expériences. Pour comprendre une situation déjà vécue, un script est sélectionné et activé. Lors d'une situation nouvelle, un script n'est pas appris *ex nihilo* ; un nouveau script est construit à partir d'un autre en l'adaptant à la situation nouvelle. Les scripts peuvent aussi être enrichis et adaptés lors de leur utilisation. Ils ont ainsi mené à la théorie de la mémoire dynamique selon laquelle les processus cognitifs de compréhension, de mémorisation et d'apprentissage utilisent la même structure de mémoire. Cette structure, dénommée "*Memory Organization Packets*"(MOP), contient les descriptions d'expériences passées et de situations stéréotypées. La première implantation de système de RÀPC, développée par J. Kolodner, est basée sur cette structure [50].

2.2.1 Le modèle

Le raisonnement à partir de cas peut se voir comme une fonction **RÀPC,** qui, à l'énoncé d'un nouveau problème ou problème *cible* `P`, fait correspondre une solution `Sol(P)`, en tirant parti d'un ensemble de *cas*, qui sont des problèmes déjà résolus accompagnés de leur solution. Un cas mémorisé, ou cas *source*, est la donnée d'un couple énoncé de problème – solution `(P,Sol(P))` ; une *base de cas* est un ensemble fini de cas `Base-de-cas` $= \{(P_i, Sol(P_i))/i = 1, .., n\}$. Il est classique de considérer le problème cible comme un nouveau cas `cas-cible` $= (cible, \perp)$ – où $\perp$ correspond à l'absence de solution – pour lequel il faut construire `Sol(cible)` en s'aidant d'un cas mémorisé `cas-source` $= (source, Sol(source))$.

L'implantation de ce mode de raisonnement a conduit à la notion de système de RÀPC : pour résoudre un problème, qu'il soit de conception, d'interprétation, de diagnostic, de planification, etc., un système de RÀPC s'appuie sur des connaissances d'ordre général et des connaissances particulières aux problèmes manipulés. Ces connaissances sont mobilisées au cours d'un cycle, associées à la fonction **RÀPC,** composé classiquement de cinq opérations [1] : élaboration, remémoration, adaptation, révision et mémorisation comme présenté figure 2.3 et détaillé ci-après.

Les connaissances générales sur le domaine étudié sont enregistrées dans ce que l'on nomme une théorie du domaine. Elle sont mobilisées notamment lors de l'opération de remémoration. Elles permettent d'établir les relations de similarité entre un problème cible et des problèmes sources de la base de cas. Le cas représente des connaissances particulières propres à un problème auquel est associée sa solution. Les connaissances embarquées dans les cas sont alors utilisées pour permettre l'adaptation d'une solution d'un cas source pour résoudre un problème cible.

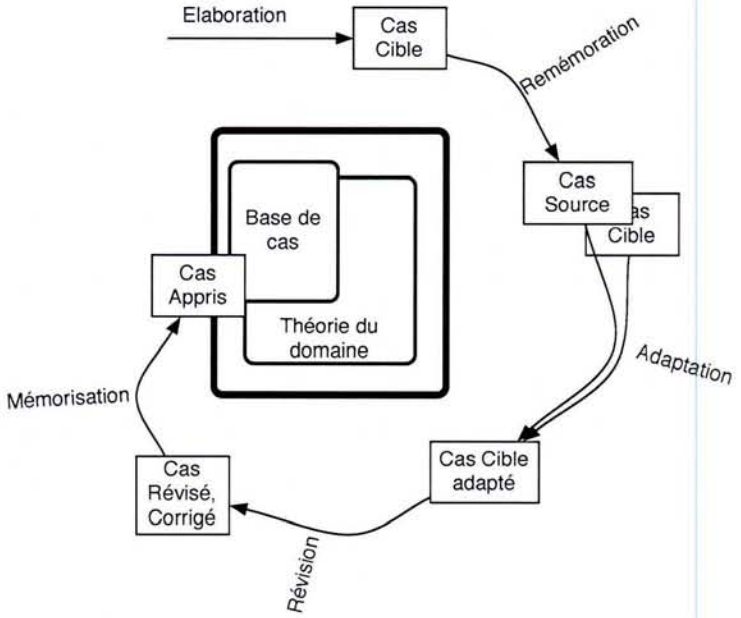


FIG. 2.3 – Les étapes du raisonnement à partir de cas, d'après [1].

2.2.2 Le cycle du RàPC

2.2.2.1 Elaboration

Dans le cas général, lorsqu'un problème est posé, les connaissances à son sujet peuvent être de nature très différentes. Certaines de ces connaissances peuvent être totalement étrangères à la résolution du problème. D'autres, en revanche, peuvent être absentes et pourtant indispensables pour permettre d'envisager une résolution. La première phase du raisonnement a donc pour but de compléter et/ou de filtrer la description du problème en se fondant sur des connaissances générales et notamment sur les connaissances d'adaptation qui seront mises en œuvre lors de la phase d'adaptation. Cette première phase de caractérisation du problème, de sa description, correspond à la phase d'élaboration. Elle permet l'affectation au nouveau cas cible de descripteurs possédant une sémantique liée au problème. Cette affectation de descripteurs doit être envisagée dans un but d'anticiper au maximum l'adaptabilité des cas qui seront remémorés.

2.2.2.2 Remémoration

Une fois le problème clairement posé, élaboré, la phase de remémoration commence. Elle consiste à retrouver, dans la base de cas, des cas similaires dont la solution enregistrée pourra être adaptée. La similarité est déterminée par le degré d'appariement entre le problème cible et les problèmes sources de la base de cas. La remémoration a pour but de trouver, dans une base de cas, un ou des cas dont le problème soit similaire au problème cible. Le choix d'une mesure de similarité est déterminant pour l'efficacité du RÀPC. En particulier, la sémantique de la mesure de similarité choisie doit être liée à la tâche de

résolution du problème posé [93]. Lieber [72] propose de prendre en compte le processus d'adaptation au moment de la remémoration. Il définit la notion de *chemin de similarité* fondée sur une relation de généralité entre problèmes. Le chemin de similarité est construit lors de la mise en correspondance des problèmes, pour être réutilisé lors de la phase suivante d'adaptation. On parle alors de remémoration guidée par l'adaptation. Cette approche garantit de pouvoir adapter les cas remémorés. Elle limite la recherche au cours de la remémoration et avance le travail pour l'adaptation. Les deux opérations ne sont plus considérées indépendamment.

2.2.2.3 Adaptation

L'adaptation a pour tâche de construire une solution `Sol(cible)` du problème cible en s'appuyant sur la solution `Sol(source)` du cas remémoré, appelé cas source et noté `(source, Sol(source))`. Dans le cadre du RÀPC, le "carré d'analogie" peut être spécifié comme présenté figure 2.2 pour représenter la tâche à réaliser par la fonction d'adaptation.

L'adaptation a fait l'objet de nombreuses propositions dans les premières années de la recherche sur le RÀPC [93]. Mais cette opération délicate est rarement menée à son terme dans les applications concrètes. En effet, si les méthodes à appliquer pour la remémoration sont aisément imaginables, les méthodes d'adaptation paraissent plus multiformes. Trois voies principales de recherches peuvent être définies [73] :

- Les démarches unificatrices. Elle ont pour but de proposer des principes, des algorithmes pouvant aider le concepteur d'un système de RÀPC à mieux appréhender l'adaptation dans sa globalité. C'est notamment dans ce cadre que se situent les travaux sur les chemins de similarité. Un chemin de similarité peut être utilisé pour construire un *chemin de modification* [86] servant au calcul des adaptations à apporter sur le cas source pour proposer une solution au cas cible.
- Les méthodes d'acquisition de connaissances d'adaptation. Ces méthodes cherchent à modéliser et à implanter des connaissances d'experts du domaine. Elle font appel aux techniques d'ingénierie des connaissances. C'est la méthode que nous avons privilégiée sans toutefois la mener à terme.
- La recherche de catalogues de stratégies d'adaptation. Le but est d'identifier les différentes pratiques d'adaptation afin de proposer des solutions génériques en fonction des situations observées.

2.2.2.4 Révision et mémorisation

Un fois une adaptation trouvée, la solution est présentée à l'utilisateur du système. Cette solution peut alors être testée. Si elle ne convient pas, il est possible de renseigner le système sur les causes de l'échec. Le système doit alors réviser ses connaissances sur le cas source ayant servi de base pour la résolution du problème et/ou sur les adaptations qui ont été effectuées. Il est même possible qu'il faille augmenter la théorie du domaine pour rendre compte de la spécificité du nouveau cas par rapport aux cas de la base. Ce mécanisme de révision a notamment été étudié dans le cadre du développement du système de RÀPC précurseur qu'est CHEF [113]. Si le nouveau cas est digne d'intérêt, il peut alors être enregistré pour enrichir la base de cas. Ces dernières phases du raisonne-

ment sont généralement à la charge de l'expert du domaine responsable du système. On peut alors considérer que la fonction d'apprentissage qui consiste à ajouter des nouveaux cas ou modifier des connaissances pour résoudre des situations d'échec correspond à un apprentissage supervisé.

2.3 Le RàPC en milieu hiérarchique

Dans le cadre de nos travaux, nous avons décrit nos connaissances sur les cas à l'aide du système de représentation en logiques de descriptions RACER. Cette approche nous offre un environnement où l'organisation des connaissances est faite dans un cadre hiérarchique. Aussi, la classification est-elle à la base de notre implantation du raisonnement à partir de cas. Nous présentons ici, sur la base de [76], en quoi les opérations de remémoration et d'adaptation peuvent tirer avantage de cette situation.

2.3.1 L'indexation

Une base de cas peut éventuellement être munie d'une relation d'ordre, qui permet de comparer les cas entre eux et de déterminer leur *degré de généralité*. Toutefois, ce n'est généralement pas sur les cas eux-mêmes, ni sur l'énoncé des problèmes qu'est définie une telle relation d'ordre, mais plutôt sur une *généralisation* de cet énoncé, appelée *index* : pour un cas source K = `(source,Sol(source))`, l'index de K, noté `idx(source)`, est un énoncé de problème "plus général que" `source`. L'index d'un cas est censé décrire les caractéristiques essentielles qui sont associées à l'énoncé du problème et à la construction de sa solution.

Intuitivement, `idx(source)` matérialise la partie de `source` qui joue directement un rôle dans le processus de construction de `Sol(source)`, ce qui s'accorde avec la notion d'*empreinte* — *footprints* — telle qu'elle est introduite dans [122]. L'indexation de $K = (source, Sol(source))$ produit `idx(source)` et consiste à enlever dans `source` les éléments qui ne jouent pas de rôle dans `Sol(source)`. Si l'on sait décider qu'un élément de `source` joue ou ne joue pas de rôle dans `Sol(source)`, alors le processus d'indexation peut être automatique — comme dans [122] et [49] — mais il reste bien souvent manuel et dépendant de connaissances propres au domaine d'étude, voire de spécialistes, comme expliqué dans [72].

Il faut faire ici deux remarques :

R_1 : La première est qu'un index `idx(source)` est un énoncé de problème au même titre que l'énoncé `source` lui-même. En particulier, les énoncés de problèmes et les index sont supposés être tous représentés dans le même formalisme.

R_2 : La seconde est que l'indexation ne se conçoit qu'en présence d'une solution : pour un cas K = `(source,Sol(source))`, il est supposé que la solution `Sol(source)` est généralisable en une solution de `idx(source)`. En particulier, pour un problème cible $(cible, \perp)$, `idx(cible)` est identique à `cible` : puisque `Sol(cible)` n'est pas connue, c'est `cible` dans son ensemble qui est susceptible de jouer un rôle dans le processus de résolution du problème cible ; autrement dit, l'index de `cible` est `cible` elle-même.

2.3.2 La remémoration

L'opération de remémoration consiste à retrouver un cas `cas-source` *similaire* au cas étudié `cas-cible`. La recherche du cas mémorisé ne se fait pas sur la donnée du cas cible, mais sur la donnée des index associés à `cas-source` et `cas-cible` : c'est grâce aux index qu'est évaluée la similarité entre `cas-source` et `cas-cible`. L'opération de remémoration est ici considérée comme une opération de classification de `cible` dans une hiérarchie notée $\mathcal{H}_{idx}$ et appelée *hiérarchie des index*. La hiérarchie $\mathcal{H}_{idx}$ contient deux types de classes : d'une part, des classes qui représentent des énoncés de problèmes qui sont des index et qui permettent donc d'accéder à des cas sources `(source,Sol(source))` ; d'autre part, des classes *abstraites* qui ne sont reliées à aucun cas source mais dont le rôle est de factoriser les informations présentes dans les index afin de favoriser le processus de classification.

Plus précisément, l'opération de remémoration se divise en deux étapes :

- Construction de la représentation de la cible, qui peut être vue comme une "indexation triviale" puisque $idx(cible) =$ `cible` ; cette construction est à rapprocher de la phase d'*élaboration* introduite dans [36].
- Classification : `cible` est classifiée dans $\mathcal{H}_{idx}$; tous les énoncés de problèmes sources dont l'index est un subsumant de `cible` fournissent une solution qui peut être potentiellement réutilisée pour produire une solution du problème cible `cible`.

La classification de `cible` dans $\mathcal{H}_{idx}$ peut déterminer plusieurs cas sources qui peuvent être réutilisés pour résoudre le problème cible. Si un ou plusieurs cas sources sont remémorés, alors un cas est sélectionné et l'opération d'adaptation est activée, ce qui correspond au processus de *classification dure*. En revanche, si aucun cas source n'est remémoré — les subsumants les plus spécifiques de `cible` dans $\mathcal{H}_{idx}$ sont des classes abstraites, $\top$ en particulier — alors le processus de *classification élastique* est activé.

2.3.3 La classification dure, les chemins de similarité et l'adaptation

L'opération d'adaptation se combine avec l'opération de remémoration par l'intermédiaire de la notion de *chemin de similarité* qui fait référence à une suite de relations liant l'énoncé du problème source `source` à l'énoncé du problème cible `cible`.

Plus précisément, le rôle du processus de classification dure est de mettre en évidence des chemins de similarité $sim(source, cible)$ de la forme :

$$source \sqsubseteq idx(source) \sqsupseteq cible \qquad (1)$$

La relation $source \sqsubseteq idx(source)$ indique que l'index de l'énoncé du problème `source` est plus général que l'énoncé `source`, ce qui est le propre d'un index. La relation $idx(source) \sqsupseteq cible$ indique que l'index de l'énoncé du problème `source` est plus général que l'index de l'énoncé du problème cible, qui est `cible` elle-même.

Voici un exemple, dans le monde des blocs, d'un chemin de similarité qui peut être obtenu après classification dure :

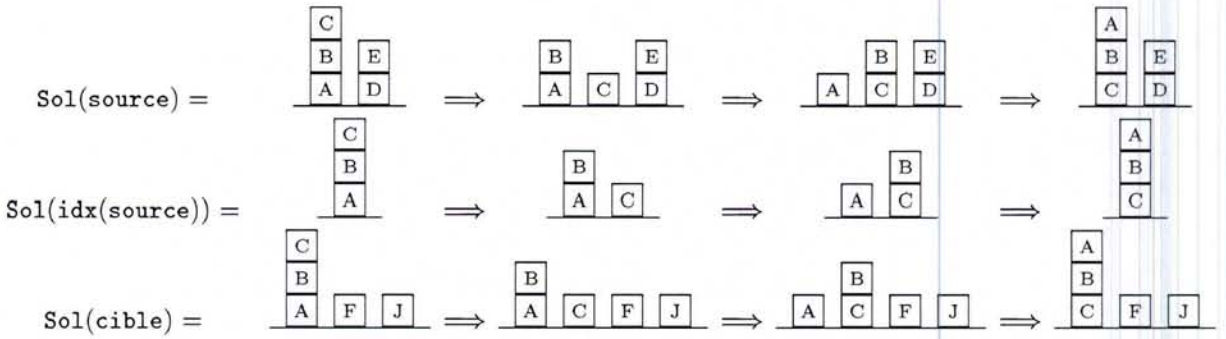
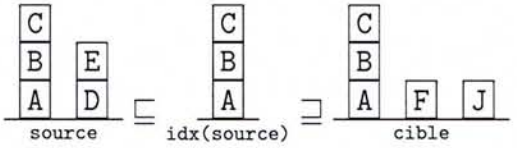


FIG. 2.4 – Un exemple d'adaptation à l'aide d'un chemin de similarité reposant sur une classification dure. La double flèche $\Longrightarrow$ indique les transitions entre les différents états du processus de résolution. On sait comment inverser A et C dans source, on adapte pour le faire dans cible.



Lorsqu'un chemin de similarité de la forme `(1)` est mis en évidence, l'opération d'adaptation s'appuie essentiellement sur deux étapes, la *généralisation* et la *spécialisation*. Une première étape dans l'adaptation consiste à généraliser la solution `Sol(source)` associée au cas source pour produire une solution généralisée $Sol(idx(source))$ (c'est l'objet de la remarque R_2 du paragraphe 2.3.1). Une seconde étape consiste à spécialiser `Sol(idx(source))` en `Sol(cible)`.

Un exemple d'adaptation dans le monde des blocs est donné à la figure 2.4. Cet exemple va nous servir aussi à expliciter la relation qui existe entre l'énoncé `source` et l'index `idx(source)`. Chaque ligne du schéma correspond à la résolution d'un problème avec ses différentes étapes. La ligne du haut montre la solution associée au cas source, qui est généralisée en la ligne du milieu, qui correspond au problème associé à l'index du cas source. Enfin, le problème cible correspond à la ligne du bas et la solution obtenue pour le cas cible est une spécialisation de la ligne du milieu. L'énoncé `source` comprend deux piles de blocs, CBA et ED de haut en bas ; l'index de cet énoncé est la seule pile CBA. En effet, la solution proposée ici consiste à transformer la pile CBA en une pile ABC, sans que la pile ED ne soit modifiée. Ainsi, l'index `idx(source)` ne tient compte que de la seule pile CBA, et le processus de résolution du problème `idx(source)` ne comporte que les éléments nécessaires à la résolution du problème `source`.

Le schéma de la figure 2.4 montre les deux étapes de généralisation – spécialisation permettant de trouver la solution d'un problème cible après classification dure. Ainsi, dès lors que le processus de classification dure a établi un chemin de similarité entre un énoncé `source` auquel est associée une solution `Sol(source)` et un énoncé `cible`, une suite d'opérations de généralisation et de spécialisation permet d'adapter la solution `Sol(source)` en une solution `Sol(cible)`.

2.3.4 La classification élastique

Le processus de classification dure s'appuie sur un appariement *exact*. Cependant, il arrive que les caractéristiques des énoncés des problèmes associés aux cas de la base ne s'apparient pas directement avec les caractéristiques de l'énoncé du problème cible. En particulier, l'équation `(1)` n'est plus vérifiée et les seuls subsumants de `cible` dans $\mathcal{H}_{idx}$ sont des classes abstraites ($\top$ en particulier). C'est alors qu'intervient le processus de *classification élastique* dont le principe est le suivant : "déformer" `source` et `cible` de telle sorte que l'équation `(1)` soit vérifiée par les énoncés déformés de `source` et `cible`. En d'autres termes :

$$\text{classification élastique} = \text{classification dure} + \text{déformations.}$$

Plus précisément, étant donné un ensemble de fonctions de déformation, le processus de classification élastique cherche à construire un chemin de similarité $sim(source, cible)$ de la forme :

$$source \sqsubseteq idx(source) \simeq \varphi(idx(source)) \sqsupseteq \psi(cible) \simeq cible \quad (2)$$

Les fonctions de déformation élémentaires dont il est question ici sont des fonctions de généralisation, de spécialisation, de changement de type et de modification qui dépendent du domaine où le problème est posé. Dans l'équation (2), φ et ψ sont des fonctions de déformation qui sont supposées vérifier les propriétés suivantes :

R_3 : si une solution `Sol(idx(source))` est connue, alors on peut construire une solution $Sol(\varphi(idx(source)))$,

R_4 : si une solution $Sol(\psi(cible))$ est connue, alors on peut construire une solution `Sol(cible)`.

Les propriétés R_3 et R_4 donnent le sens de la relation $\simeq$ dans l'équation `(2)` : il est possible de passer de l'énoncé du problème `idx(source)` à l'énoncé $\varphi(idx(source))$ en conservant le fait que le problème considéré possède une solution ; il en va de même pour $\psi(cible)$ et `cible`. Les propriétés R_3 et R_4 sont importantes et jouent un rôle lors de l'opération d'adaptation, du même ordre que la généralisation d'une solution de `source` à `idx(source)`, comme indiqué dans la remarque R_2 du paragraphe 2.3.1. Ainsi, étant donné un chemin de similarité de la forme `(2)`, l'opération d'adaptation s'appuie sur trois opérations, la généralisation et la spécialisation, comme lors de l'adaptation d'un chemin de similarité de la forme `(1)`, avec en plus la prise en compte des déformations régies par les fonctions φ et ψ. L'adaptabilité des étapes de déformation repose alors sur les propriétés R_3 et R_4.

Un exemple d'adaptation associé à une classification élastique est donné à la figure 2.5 qui montre un chemin de similarité de la forme `(2)`, obtenu après classification élastique, dans le monde des blocs. Dans cet exemple, l'application de la fonction φ consiste à passer du bloc étiqueté C à un bloc générique étiqueté par la variable x : l'application de φ correspond donc à une généralisation d'étiquette ; x est une variable et le bloc dénoté par x peut être substitué par tout autre bloc. L'application de la fonction ψ consiste à déplacer le bloc J sur le bloc B, et correspond à un déplacement élémentaire de bloc.

Les fonctions de déformation φ et ψ et les adaptations correspondantes, régies par R_3 et R_4, dépendent du domaine d'application. Les fonctions qui ont été évoquées ici

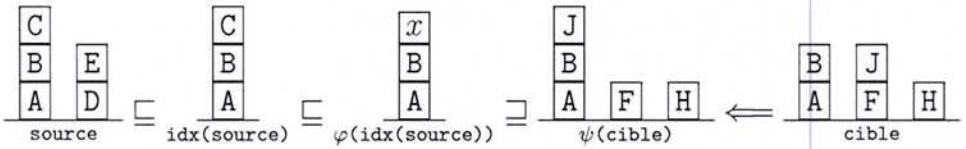


FIG. 2.5 – Un exemple de chemin de similarité obtenu après classification élastique. Le bloc étiqueté par x correspond au bloc générique qui peut être substitué par tout autre bloc.

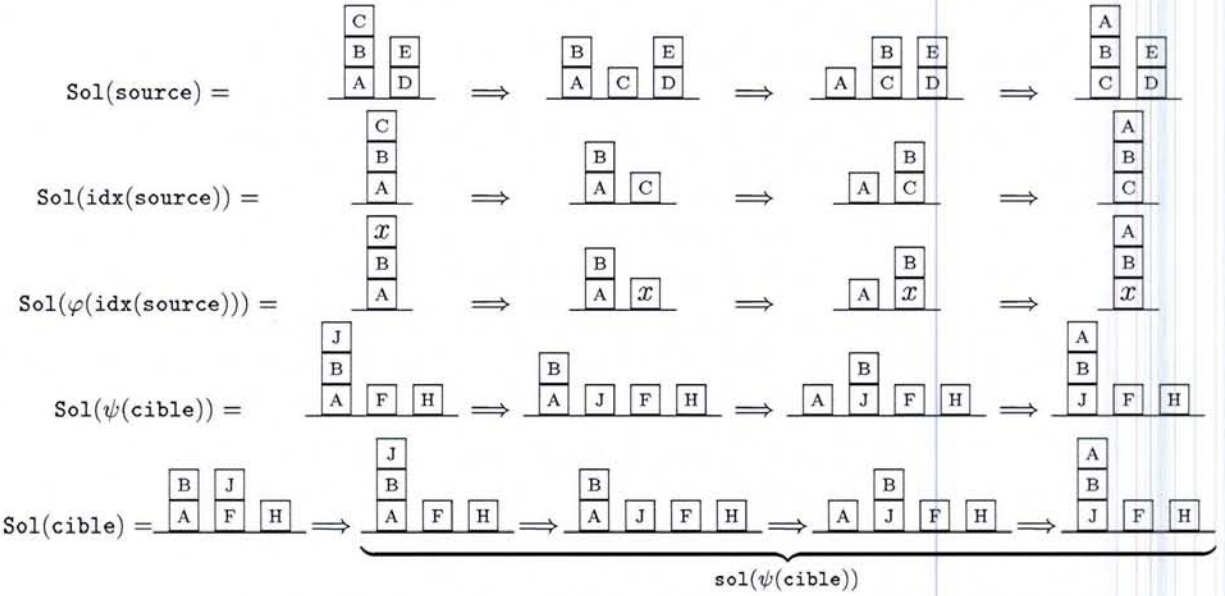


FIG. 2.6 – Un exemple d'adaptation à l'aide d'un chemin de similarité reposant sur une classification élastique.

— à savoir les généralisations pour φ et les transformations d'états (déplacement d'un bloc) pour ψ — se retrouvent dans de nombreuses circonstances : les généralisations φ s'utilisent dès lors qu'il existe une relation de généralité entre énoncés de problèmes ; les transformations d'états ψ s'utilisent couramment en planification.

Le processus d'adaptation associé au chemin de similarité de la figure 2.5 est donné à la figure 2.6. Le problème cible consiste à passer de la configuration BA JF H à la configuration ABJ F H. La résolution du problème `cible` s'appuie sur la solution `Sol(source)` associée à `source`, *via* l'application des fonctions de déformation φ et ψ.

Un autre exemple détaillé d'adaptation associée à une classification élastique est présenté dans [72].

Les fonctions φ et ψ peuvent être vues comme des compositions de fonctions élémentaires auxquelles un coût est associé. Le coût du chemin de similarité est calculé comme la somme de ses coûts élémentaires. On obtient ainsi une évaluation de la similarité entre cible et source.

2.3.5 Les difficultés de l'adaptation

Le processus de remémoration s'appuyant sur les classifications dure et élastique peut être qualifié de *remémoration guidée par l'adaptabilité*, comme présenté dans [119] et [118]. Il met en valeur la similarité entre deux énoncés de problèmes, mais aussi "en quoi" les deux énoncés sont similaires et comment "passer" de l'un à l'autre par l'intermédiaire d'un chemin de similarité. Le meilleur cas source devant être adapté pour résoudre le problème cible est alors celui qui réalise le chemin de similarité de coût minimal, un coût élémentaire étant associé à chaque opération du chemin de similarité (généralisation, spécialisation et déformation).

L'*adaptabilité* est une notion purement *syntaxique* qui ne dépend que de la capacité du système à proposer une solution par adaptation — sans préjuger de la correction de la solution — et qui s'oppose à la *validité sémantique* de l'adaptation, qui est la capacité du système à fournir une solution correcte. Ainsi, le problème de la validité de la solution obtenue par adaptation se pose dans les mêmes termes que pour le raisonnement par analogie [18]. Pour notre part, nous considérons qu'il ne peut être réglé qu'en interaction avec l'extérieur.

2.4 Le raisonnement à partir d'explications

Le raisonnement à partir d'explications est un cas particulier de raisonnement à partir de cas développé par R.C. Schank alors qu'il cherchait à définir l'intelligence. Il pose deux questions auxquelles il cherche à répondre :

- Quels sont les critères à prendre en compte pour déterminer l'intelligence ou la pensée ?
- Que signifie comprendre ?

Schank pense que c'est la notion d'explication qui est au cœur de l'intelligence. Ce sont les explications que nous pouvons produire qui donnent du sens au monde qui nous entoure [117]. Le sens du monde prend naissance dans nos expériences de tous les jours qui font à chaque fois le voir différemment. Des valeurs sociales ou personnelles différentes donnent des perceptions différentes.

Comprendre une personne ou un animal est différent de comprendre ce que nous avons lu, entendu ou vu. Notre connaissance, notre compréhension évolue avec le temps, avec nos expériences : nous apprenons continuellement ! Comprendre, c'est être capable de dire des choses justes, d'avoir des émotions, de faire des déductions surprenante mais justes. Comprendre, c'est être capable d'expliquer ses actions aux autres et à soi même (le plus difficile à atteindre). Par exemple MYCIN, dans le monde des blocks, peut expliquer ses actions pour arriver au résultat demandé. Comprendre le monde, c'est poser des questions et apporter des réponses. Ce ne sont pas les tâches exécutées qui sont importantes (porteuses d'intelligence) mais bien leur choix et leur ordre d'exécution en expliquant pourquoi.

Schank s'intéresse au début de ses recherches aux problèmes de linguistique et au traitement automatique du langage naturel. À contre-courant des travaux sur les syntaxes et grammaires (Chomsky notamment), Schank propose une approche essentielle-

ment conceptuelle de la compréhension du langage naturel. « Oublier les mots au profit des concepts... ». Dans le courant des années 70, le concept de *script* est proposé pour décrire des comportements stéréotypiques. Un script devenu célèbre est l'exemple du restaurant : *entrer, s'asseoir, obtenir la carte, lire le menu, commander, être servi, manger, payer, sortir* [117]. Pour manger au restaurant, il vous suffit de connaître le script décrivant les différentes actions, les différentes phases, et de le jouer. Les scripts seraient des descriptions très personnelles des différentes situations de la vie courante, fruits de l'apprentissage au cours de nos expériences. Mais toutes nos actions ne peuvent se résumer à la recherche et à l'activation du bon script dans la mémoire personnelle. Lors d'une situation nouvelle, nous n'apprenons pas totalement un nouveau script. Au contraire, nous partons de situations connues pour dériver un nouveau script en l'adaptant à la nouvelle situation. Ses travaux sur la mémoire dynamique et l'apprentissage, pour l'étude de la compréhension d'histoires énoncées en langage naturel, ont mené Schank à proposer un modèle d'organisation de la mémoire sous forme d'une structure de mémoire dynamique pour la recherche d'informations : les MOPs (Memory Organization Packages) [116].

Un MOP représente une unité de base de la mémoire dynamique. Dans les systèmes fondés sur ce modèle de mémoire, un MOP représente une structure servant à la description d'un cas qui en est une instance. Un MOP permet de décrire la connaissance disponible dans une classe d'événements et particulièrement les événements complexes. Il contient une liste de "traits normaux" représentant les caractéristiques de base des événements qu'il modélise. Les traits normaux décrivent le type d'événement, les buts associés au MOP, les acteurs impliqués, etc. Ils représentent en quelque sorte la description d'un script. Mais contrairement aux scripts, qui sont des structures statiques, les MOPs sont organisés en réseaux interconnectés et en hiérarchie de spécialisation. Un MOP donné peut être lié à d'autres MOPs décrivant des sous-événements ; il peut être spécialisé en un MOP plus spécifique. Un MOP représente un concept d'événement pouvant être instancié pour décrire un événement particulier. Un MOP peut contenir des exemples de liens vers des instances particulières. Enfin, un MOP peut inclure la description d'exceptions par l'utilisation de liens vers des instances décrivant des résultats non prédits par ce MOP.

Pour expliquer une entrée (cible), on s'appuie sur une mémoire structurée (les MOPs) où placer (classer) cette cible. Les MOPs constituent des sources d'actions standard. Comprendre la cible, c'est expliquer en quoi elle décrit quelque chose de connu. Sans classification de la cible, en l'absence d'un MOP préparé pour la reconnaître, une demande d'explication nouvelle doit être engagée. Pour commencer la résolution d'un problème, Schank propose de faire référence à un ensemble de questions standard relativement limitées les *Explanation Questions* (EQ). Elles font référence à trois types d'explications attendues :

- A - Type de défauts
- B - Type d'événements
- C - Type de buts

Chaque EQ est aussi un index à des schémas d'explication, *Explanation Pattern*, (XP) qui peuvent être adapté pour produire de nouvelles explications. Les XP représentent des explications fossilisées. Ils fonctionnent un peu comme les scripts d'événement. Quand un XP est activé, il connecte un événement à expliquer avec une explication qui a été enregistrée au cours de l'explication d'un ancien événement similaire à l'événement présent.

L'utilisation classique d'un XP est la suivante :

1. Identification d'une séquence d'événement qui nécessite une explication.
2. Etablir un index des questions qui peuvent éventuellement mener à un XP.
3. Rechercher l'XP pertinent attaché à cet index
4. Remplacer l'XP dans la séquence d'événement.

Mais une autre utilisation est souvent envisagée. Cela se passe lorsqu'il n'y a pas d'XP en mémoire qui convienne pour expliquer une situation. Dans ce cas, l'on peut utiliser les explications retrouvées, mais seulement comme point de départ à la création d'une nouvelle explication.

2.5 Conclusion

Le raisonnement à partir de cas peut être utilisé dans différents contextes notamment pour des applications :

- De résolution de problèmes de classification, de diagnostic, de configuration, de design et de planification.
- De conception de système d'aide à la décision ("decision support"). Actuellement les systèmes de "help-desk" sont la principale application commerciale du RÀPC. Ce sont des systèmes aviseurs qui fournissent aux décideurs/analystes des recommandations sur les actions à entreprendre dans des situation routinières.
- De préservation et d'exploitation de la connaissance des entreprises. Le domaine de la gestion de connaissances est actuellement en pleine expansion.
- De type "recherche d'information" sur des domaines restreints comme par exemple l'exploitation de documents de type "*Frequently-Asked Questions*" (FAQ : paires de questions-réponses).

L'approche du raisonnement à partir de cas offre de nombreux avantages. Elle permet d'éviter, dans une certaine mesure, les problèmes d'acquisition de connaissances qui rendent difficile la construction de bases de connaissances de taille importante. Pour certaines applications, l'approche RÀPC est plus simple à mettre en œuvre que les approches basées sur un modèle du domaine. Elle permet de construire rapidement des prototypes qui pourront être améliorés avec l'enrichissement de nouveaux cas. De plus, le RÀPC est particulièrement bien adapté pour les applications ayant les caractéristiques suivantes :

- La tâche est accomplie par des humains expérimentés dans leur domaine et ces expériences sont disponibles dans une base de données, dans des documents ou chez un expert humain ;
- Une analyse détaillée du domaine n'est pas nécessaire pour obtenir des solutions satisfaisantes et la tâche n'exige pas une solution optimale ;
- Un modèle du domaine ne peut être élaboré parce que le domaine est mal formalisé (peu de documentation, expert non disponible) ou parce qu'il n'existe pas de principes généraux qui sont éprouvés (par exemple, comment investir à la bourse) ;
- Les situations sont répétitives et les solutions sont réutilisables. Ces situations, dites monotones, sont telles que de petites différences dans le problème entraînent de

petites différences dans la solution. De plus, une solution valide à un moment le demeure à un autre moment.

S.C.D. - U.H.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

Chapitre 3

La modélisation des organisations spatiales agricoles

Sommaire

3.1 Problématique agronomique 46
3.1.1 Transformation des espaces ruraux 46
3.1.2 Production agricole et gestion de l'environnement 48
3.2 Modéliser l'organisation spatiale des exploitations agricoles 51
3.2.1 L'analyse des exploitations agricoles 51
3.2.2 La méthode des chorèmes 52
3.2.3 L'approche spatiale des exploitations agricoles 54
3.3 Exemple d'application de la méthodologie : le cas de l'étude du Sauveterre 59
3.3.1 Étape préparatoire à l'enquête 59
3.3.2 L'enquête en exploitation 60
3.3.3 Le chorème d'exploitation 64
3.4 Questions soulevées par la méthodologie appliquée 64

Ce chapitre présente le positionnement des agronomes et les questions qu'ils posent aux informaticiens pour modéliser des organisations spatiales agricoles. Le contexte agricole a changé, les espaces ruraux se transforment. La société attend des agriculteurs qu'ils répondent à de nouvelles fonctions dans les territoires. Cela renouvelle les approches des agronomes qui doivent prendre en compte le territoire et s'intéresser à la dimension spatiale des activités agricoles.

Après avoir re-situé ces problématiques dans leur contexte historique, nous présentons comment les agronomes avec lesquels nous travaillons dans le cadre de cette thèse, ont développé leurs recherches, notamment à travers deux chantiers en Lorraine d'une part et sur les Grands Causses d'autre part (3.1).

Pour étudier le fonctionnement des exploitations et comprendre leur évolution, les agronomes ont développé une approche de modélisation spatiale qui formalise les principes organisateurs de l'espace et représente les exploitations agricoles sous forme de chorèmes, en s'appuyant sur des travaux en géographie. Ils développent une méthodologie d'enquête au sein des exploitations agricoles pour modéliser l'organisation spatiale des exploitations agricoles (3.2).

Un exemple d'utilisation de cette méthode est étudié (3.3). La problématique concerne la maîtrise de l'embroussaillement sur le causse de Sauveterre. Les questions soulevées par l'utilisation de cette méthode sont exposées. Elle ont mené à la démarche de formalisation de l'approche spatiale qui est l'objet du chapitre suivant.

3.1 Problématique agronomique

Les transformations, en cours dans les espaces ruraux, induisent de nouvelles attentes vis à vis de l'agriculture et renouvellent les approches agronomiques, en particulier au niveau des exploitations agricoles.

3.1.1 Transformation des espaces ruraux

Au lendemain de la seconde guerre mondiale, la volonté affirmée en Europe d'une autosuffisance alimentaire a été le moteur d'un bouleversement du monde rural. Un modèle productiviste a été instauré, associant une modernisation technique et une restructuration des exploitations agricoles. Le but de ce modèle, soutenu par la première version de la PAC[9], était d'accroître rapidement les volumes produits par l'agriculture. Une des conséquences de ce modèle fut une réduction importante du nombre des exploitations, accompagnée d'une accélération de l'exode rural initié depuis la révolution industrielle du XIXème siècle.

Ce modèle a produit des résultats au-delà des espérances si bien que dès les années 75-80, une surproduction structurelle s'est installée. Durant des années, cette surproduction a alimenté les exportations agro-alimentaires de la France. Mais ce modèle, engendrant une augmentation constante des volumes produits, est largement contesté depuis 15 à 20 ans. L'intervention directe sur les marchés, hier forme essentielle d'aide publique à l'agriculture,

[9]Politique Agricole Commune.

est en contradiction avec les règles de l'OMC[10]. Dans le même temps, les attentes de la société ont évolué. L'exigence de production de nourriture est toujours présente, mais fortement conditionnée par des exigences nouvelles de sécurité et de qualité. Elles se manifestent aussi vis-à-vis d'autres fonctions :

- Gestion de l'environnement par la maîtrise des impacts sur les ressources naturelles, sur la bio-diversité,
- Sociale avec une volonté du maintien d'un tissu rural et d'emplois,
- Culturelle et symbolique : sauvegarde des cultures alimentaires, entretien du rapport des sociétés à la nature et au vivant à travers les activités agricoles, gestion des paysages, etc.

Ces attentes de la société se traduisent par des dynamiques d'évolution propres au monde rural. Aujourd'hui, les campagnes connaissent un retour de population après avoir vécu une longue période d'exode rural [5]. Ce nouveau solde migratoire positif s'accompagne d'un changement dans la composition socio-professionnelle du monde rural. Ce qui lui donne de plus en plus un statut d'espace à fonction résidentielle et/ou récréative. Cette évolution se traduit par une diminution du nombre d'agriculteurs investis dans les collectivités territoriales (mairies, conseils généraux, etc.).

Parallèlement, les politiques liées au monde rural évoluent également. Les lois de décentralisation ont délégué aux collectivités territoriales les compétences de gestion de l'espace rural. En 1995 est créé le Fonds de Gestion de l'Espace Rural pour pouvoir soutenir les initiatives des acteurs locaux. Aujourd'hui, les nouvelles lois d'orientation agricole et d'aménagement et de développement durable du territoire insistent sur la mise en place d'outils de gestion à l'échelle locale (Contrats Territoriaux d'Exploitation, Pays, etc.). De ce fait, les pouvoirs décisionnels en matière de gestion de l'espace se déplacent vers l'échelon local.

Ainsi, les nouvelles fonctions de l'espace rural sont de plus en plus prises en compte dans les objectifs de gestion. Il est demandé aux agriculteurs, encore utilisateurs majoritaires de l'espace, de prendre en compte la multifonctionnalité de cet espace et non plus seulement sa fonction productive.

Face à ces enjeux, il est important d'évaluer le caractère multifonctionnel de l'agriculture ou ses possibilités de le devenir. Le territoire rural est entretenu par les exploitants agricoles à la demande, non explicite en général, de la société toute entière, principalement citadine, et peu au fait des difficultés rencontrées par les exploitants dans cette mission. Cette redéfinition de l'agriculture, fortement liée au territoire, a pour conséquence :

- Une revalorisation des spécificités (produits du terroir),
- Une recherche de modification de la physionomie du territoire (gestion de l'espace, lutte contre l'embroussaillement, entretien du paysage...) et de ses attributs (niveau de pollution de l'eau, biodiversité...),
- Une prise en compte croissante des autres utilisations et autres utilisateurs de l'espace : contribution au développement local, complémentarités pour l'agritourisme, conflits pour le bruit, les odeurs, la pollution...

Pour accompagner ces changements, et aider à la mise en place de nouvelles règles allant dans le sens de ces nouvelles attentes tout en respectant et en développant le travail des

[10]Organisation Mondiale du Commerce.

agriculteurs, il est important d'analyser et de bien comprendre ces nouvelles dynamiques spatiales. Cela nécessite d'identifier les logiques de fonctionnement des exploitations agricoles et les modalités de gestion de l'espace, donc de décrire les caractéristiques et les utilisations du territoire des exploitations.

3.1.2 Production agricole et gestion de l'environnement

L'analyse des activités agricoles doit permettre de répondre à des questions agri-environnementales en terme de maintien de la biodiversité, d'entretien des paysages ou encore de gestion des sols [4] ou de l'eau. Ces réflexions ne sont pas déconnectées des conséquences sociales liées aux évolutions de la gestion des territoires. Qu'en est-il des relations entre les nouveaux touristes, qui profitent de cet espace pour pratiquer de nombreuses activités de plein air, et les exploitants dont les surfaces augmentent au rythme du départ des plus petits exploitants ? Quelle est l'influence, sur le travail des exploitants, des nouvelles gestions imposées par la politique agricole commune et les contrats d'entretien des paysages ? Comment s'articulent ces faits de nature et de société [71] ?

Au sein de l'INRA[11], ce sont les agronomes du département *Sciences de l'Action et Développement* (SAD anciennement nommé *Systèmes Agraires et Développement*) qui s'attachent à comprendre et à rendre compte de la diversité des fonctionnements des exploitations agricoles et de leur évolution dans ce nouveau contexte de multifonctionnalité.

Les recherches menées par les unités de Toulouse et Montpellier sur la région des Grands Causses, et de Mirecourt sur la région de Vittel constituent la base agronomique de notre travail. Ce travail s'insère par ailleurs dans le projet inter-unités *Formes d'organisation territoriales à finalités environnementales* (FORTE). Ce projet a pour but l'étude des constructions qui rendent compatibles les fonctions de production et les fonctions environnementales de l'agriculture. L'observation se fait au sein des exploitations agricoles et des dispositifs socio-économiques de gestion environnementale, sur des territoires continus, pour comprendre et maîtriser les flux biogéochimiques, les paysages produits et les biotopes.

Les deux problématiques étudiées dans le cadre de ce travail sont la gestion de la maîtrise de l'embroussaillement dans les causses et la protection de la qualité des eaux du périmètre de Vittel. Les agronomes doivent prendre en compte l'interaction entre les exploitations et leur environnement. Ils s'intéressent à l'impact de leur fonctionnement sur leur territoire à l'échelle de l'exploitation, voire au-delà, à l'échelle de la commune ou du bassin versant.

La maîtrise de l'embroussaillement et le maintien de l'ouverture des paysages est une préoccupation de nombreux partenaires sur les causses Méjan et Sauveterre. En effet, la diversification de l'utilisation du territoire, notamment avec le développement touristique de cette région, concerne aussi bien la profession agricole que les instances de gestion du Parc National des Cévennes (PNC), la DDAF [12], etc.

Depuis les années 60, les politiques agricoles et d'aménagement du territoire ont eu des effets visibles sur le paysage[13][77].

[11] Institut National de la Recherche Agronomique.

[12] Direction Départementale de l'Agriculture et de la Forêt.

[13] Programme du Ministère de l'Aménagement du Territoire et de l'Environnement (MATE) : Recréer

Les transformations du paysage dans la région des Grands Causses résultent de dynamiques anciennes, liées à des cycles de boisement/déboisement et des cycles de population/dépopulation [82, 77]. Dans un paysage marqué par le relief karstique, l'opposition entre dolines et terres labourables des fonds de vallées, dédiées à la production de céréales et de fourrages d'une part, les pentes et les plateaux dédiés à la forêt et aux parcours, d'autre part, a été entérinée depuis des siècles par le système de propriété et le cadastre.

De plus, les dynamiques de défrichage et de reconquêtes s'opposent à des dynamiques de boisement et d'enfrichement. Des boisements conséquents ont été réalisés dans les années 60, amenant actuellement à un processus de diffusion des semenciers. Par ailleurs, le bouleversement de la fonction économique et sociale des sectionnaux[14] lié à l'exode rural a dans un premier temps favorisé l'extension des accrus naturels et de l'embroussaillement. Dans un second temps, dans le cadre d'une redéfinition légale des ayants-droits, leur aménagement (défriches, éclaircies, ouverture de pistes et de points d'eau) et leur allotissement entre différents utilisateurs ont marqué le paysage.

Ces aménagements et allotissements ont conforté le processus de modernisation des exploitations agricoles des causses, en lien avec les industries de Roquefort pour la production laitière, mais aussi avec la constitution de troupeaux d'ovins viande conséquents. Toutes ces dynamiques interfèrent dans la maîtrise de l'embroussaillement [23].

Les causses Méjan et Sauveterre présentent les caractéristiques typiques de la région des Grands Causses :

- Peuplement humain réparti en domaine et hameau, plus rarement « en village », près des rares zones de champs (- de 15% de la surface sur le causse Méjan) et constituant un maillage régulier de lieux-dits habités en permanence,
- Présence ancienne d'élevage ovin valorisant ces espaces à faible productivité, élevages laitiers pour Roquefort et plus récemment pour une entreprise fromagère localisée sur le causse Méjan, et élevage viande, pour des marchés locaux ou internationaux,
- Présence également de terrains sectionnaux correspondant aux pratiques anciennes d'usages collectifs,
- Contraste entre une moitié Ouest boisée, avec une dynamique à base de pins sylvestres et une moitié Est nue, ponctuée de plantations récentes de pins noirs, avec une interrogation largement partagée sur les possibilités de maîtrise de l'embroussaillement.

La gestion de ces espaces est très largement le fait des activités agricoles, mais présente une grande diversité au niveau des exploitations agricoles et au niveau des lieux-dits. Les travaux sur le causse Méjan ont porté sur les logiques de fonctionnement des exploitations agricoles en lien avec leurs modes de mise en valeur de l'espace [60]. On observe des pratiques différenciées selon les exploitations [108]. La problématique sur le causse de Sauveterre est d'inscrire cette diversité des exploitations dans les configurations de lieux-dits [38].

L'analyse montre que la valorisation des ressources des parcours n'est pas une spécificité de la production de viande, elle n'est pas exclusive de la production intensive de lait. En

la nature, Politiques Publiques et Paysages `http://www.montpellier.inra.fr/ESR/paysages/`.

[14] Propriété indivise des habitants d'un hameau. Les sectionnaux sont des terres que se partagent les exploitants.

particulier, des troupeaux avec deux saisons de production permettent de constituer des lots à l'entretien exerçant une forte pression de pâturage ciblée. La valorisation de l'herbe des parcours passe par la combinaison, quotidienne ou saisonnière, des ressources des champs et des parcours qui a pour conséquence la constitution d'une mosaïque dans l'espace.

De plus, ni la sécurité foncière, ni la compacité des exploitations ne garantissent une gestion attentive à l'embroussaillement et aux autres usages de l'espace. Au contraire, des complémentarités s'observent dans les territoires de village, notamment en présence de biens sectionnaux et de bois, en incluant l'utilisation de lieux-dits annexes. L'observation des dynamiques en cours sur les causses Méjan et Sauveterre permet de dire que les éleveurs ont des marges de manœuvre pour maintenir un impact sur l'embroussaillement. Ces marges dépendent également de l'organisation des lieux-dits et de la façon dont les exploitations sont agencées entre elles [58].

Dans le cas du périmètre de Vittel, la Société des Eaux de Vittel s'est inquiétée de l'augmentation du taux de nitrate dans les eaux de source [28]. Suite à l'intensification de l'agriculture, les pratiques se sont totalement transformées. Les systèmes laitiers traditionnels, basés sur l'utilisation de l'herbe, sont passés à des systèmes où les cultures prennent une place grandissante pour apporter une nourriture plus riche aux vaches laitières afin d'augmenter leur production de lait. Les déplacements trop importants des animaux, pour aller aux prés, coûtent de l'énergie et réduisent d'autant la quantité de lait produite.

Pour y remédier, les exploitants ont construit de nouveaux bâtiments d'élevage, plus vastes et mieux adaptés à leur travail. Ces nouveaux bâtiments sont édifiés à l'extérieur des villages, au milieu des surfaces conservées en herbe. La majorité restante des territoires des exploitations est consacrée à la production de céréales et de maïs. Seules les surfaces difficilement cultivables sont restées en herbe. Avec le temps, les techniques culturales évoluant, certaines de ces surfaces ont pu être converties à la production de céréales. Le résultat de cette transformation est l'adoption de cultures de maïs avec l'utilisation d'engrais sur les plateaux calcaires drainants et sur les pentes les moins fortes, pour laisser l'herbe en fond de vallée dans les zones humides ou inondables. Les bâtiments d'élevage sont au centre et l'épandage du fumier lourd et non valorisé se fait à proximité sur les prairies.

Le résultat de ces pratiques conduit à une augmentation importante de la production de lait. Dans le même temps, est observée une utilisation intensive d'engrais pour les cultures associée à une faible prise en compte de l'apport de matières azotées par le fumier répandu dans le peu de prairies restantes. Il en résulte que les eaux de source se chargent en partie de ces nitrates excédentaires apportés aux cultures et aux prairies.

L'analyse des pratiques agricoles et de leur impact sur l'environnement a permis de comprendre ces phénomènes. La Société des Eaux de Vittel veut maîtriser cette augmentation des nitrates dans l'eau. Pour autant, il n'est pas question de revenir en arrière en terme de productivité ou de qualité de travail pour les exploitants. Il n'est pas non plus question de rendre désertique cette région sous prétexte que l'activité agricole peut soulever des problèmes. L'approche choisie fut d'étudier les problèmes en réunissant tous les acteurs – agriculteurs, collectivités locales, Société des Eaux de Vittel, agronomes – pour trouver des solutions efficaces dans l'intérêt de tous. La compréhension de l'interaction entre le fonctionnement des exploitations et leur environnement et notamment leur

organisation spatiale a permis de dégager certains principes d'action. Ils ont été discutés et avalisés par tous avant d'être mis en application [28].

Des recherches menées depuis plus de vingt ans sur les causses, comme de celles effectuées dans le cadre du chantier Vittel, et plus généralement des recherches menées par les agronomes du SAD, il ressort qu'il est important de prendre en compte les dimensions spatiales des activités agricoles. Des connaissances acquises au cours des années d'observation du territoire et de son évolution en parallèle à l'évolution des contextes sociaux et économiques, ces agronomes ont mis en avant que le territoire est un espace approprié et géré par les acteurs et que l'exploitation agricole représente un niveau structurant et un passage obligé du territoire.

3.2 Modéliser l'organisation spatiale des exploitations agricoles

3.2.1 L'analyse des exploitations agricoles

Pour mener leurs études, les agronomes du SAD ne s'intéressent pas uniquement à la technique agricole appliquée à la parcelle, ils s'intéressent aux interactions entre le fonctionnement des exploitations et leur territoire. Pour cela, ils étudient les pratiques des exploitants [53], en les inscrivant dans l'espace [30].

On peut cerner la transformation des activités de production et comprendre leur impact sur l'espace au niveau de l'exploitation agricole [106]. Les travaux menés ont montré qu'il existe une correspondance entre le fonctionnement des unités de production et leur structure spatiale. La structure spatiale correspond à la morphologie parcellaire et aux aspects de répartition des modes d'occupation du sol. Le fonctionnement répond à la logique de production des exploitants. Ils mettent en œuvre des pratiques contribuant également à la gestion de l'environnement (par exemple, maîtrise de l'embroussaillement et protection de la qualité des eaux).

L'organisation spatiale, vue à la fois comme un processus et comme son résultat, correspond à cette mise en relation entre structure et fonctionnement. Elle peut s'analyser par une approche spatiale et fonctionnelle, à plusieurs niveaux d'organisation [32]. Les recompositions territoriales qui résultent des fonctionnements et des évolutions sont étudiées à l'aide des concepts et outils de la modélisation spatiale [6]. On a à faire avec des phénomènes complexes, multiréférençables, pour lesquels il n'existe pas de modèles tout faits. Il faut trouver des méthodes pour interroger efficacement le terrain et rendre compte des transformations en cours.

Cette approche par une entrée spatiale que l'on relie à des fonctionnements de manière synthétique est intégratrice. Elle prolonge des approches plus fines d'autres agronomes, sur l'allotement des animaux[15], l'organisation du travail, les performances des troupeaux, les logiques de décision des agriculteurs, les stratégies d'utilisation du territoire [39]. Elle se

[15]L'allotement correspond au découpage du troupeau en différents lots, selon des caractéristiques d'âge et d'état physiologique (mise-bas, allaitement, entretien...). Au cours de l'année, des lots peuvent être regroupés ou découpés en fonction du calendrier de production et des objectifs de l'exploitant.

rattache à l'agronomie des façons de produire [27] qui ouvre sur l'agronomie des territoires [19].

3.2.2 La méthode des chorèmes

Pour rendre compte de l'organisation spatiale des exploitations, les agronomes se sont appuyés sur la méthode des chorèmes, formalisée par les géographes à un niveau régional. Ils l'ont adaptée aux questions agronomiques en développant une méthodologie d'enquête en exploitation, basée sur une approche spatiale des pratiques et synthétisée sous une forme chorématique.

La méthode des chorèmes a été mise au point par des géographes au niveau régional pour identifier les principes organisateurs de l'espace [11]. Pour Brunet (1986) :

« Un chorème est une structure élémentaire de l'espace qui se représente par un modèle graphique. Les différentes configurations spatiales, relèvent de la combinaison, éventuellement complexe, de mécanismes simples ».

Les modèles graphiques sont construits à partir des formes de base de toute représentation cartographique : point, ligne, surface, réseau et classées en sept rubriques générales : maillage, quadrillage, attraction, contact, tropisme, dynamique territoriale, hiérarchie. Ainsi par la combinaison des sept rubriques et des quatre formes de base, Brunet construit un alphabet graphique de 28 chorèmes présenté figure 3.1. L'intérêt de cet alphabet est qu'il permet de décrire, de représenter et d'interpréter des structures et des dynamiques spatiales et ce, à différentes échelles [29].

La méthode des chorèmes a semblé aux agronomes être une bonne façon de répondre à l'exigence de compréhension des dynamiques propres aux activités agricoles [29]. L'échelle où sont étudiées les activités agricoles se situe au niveau des exploitations, insérées dans de petits territoires. Les agronomes ont adapté la méthode pour en rendre compte [54]

D'un point de vue formel, un chorème d'exploitation[16] se présente comme un ensemble d'objets spatiaux, points, lignes ou surfaces, qui ont des caractéristiques morphologiques et sémantiques explicitées dans la légende, et qui ont des positions relatives rendues visuellement par le graphique. Les principes organisateurs correspondent à des modalités différentes d'affectation des surfaces ou des ressources et de circulation des animaux ou des machines entre les différentes parcelles et le siège de l'exploitation [98]. Un chorème d'exploitation est vu comme une combinaison d'entités géographiques qui ont une identité, une extension spatiale et des caractéristiques thématiques, comme dans un système d'information géographique [56].

Le chorème est censé fournir une simplification des relations entre objets, car seules celles qui sont significatives pour le modélisateur sont retenues. Une interprétation du fonctionnement de l'exploitation étudiée peut être faite, grâce à la symbolique utilisée qui renvoie à un référentiel connu du modélisateur. *« Les chorèmes présentent un indéniable intérêt comme outils pédagogiques, en raison de leur simplicité, et comme outils scientifiques, pour passer du particulier au général dans l'analyse de l'espace. Ils entendent tra-*

[16] Au sens de Brunet, le chorème correspond à la brique élémentaire, la lettre de l'alphabet, permettant la composition de la synthèse graphique. Par extension, nous utiliserons le terme de chorème d'exploitation agricole, ou composition chorématique, pour parler du schéma de l'organisation spatiale d'une exploitation agricole.

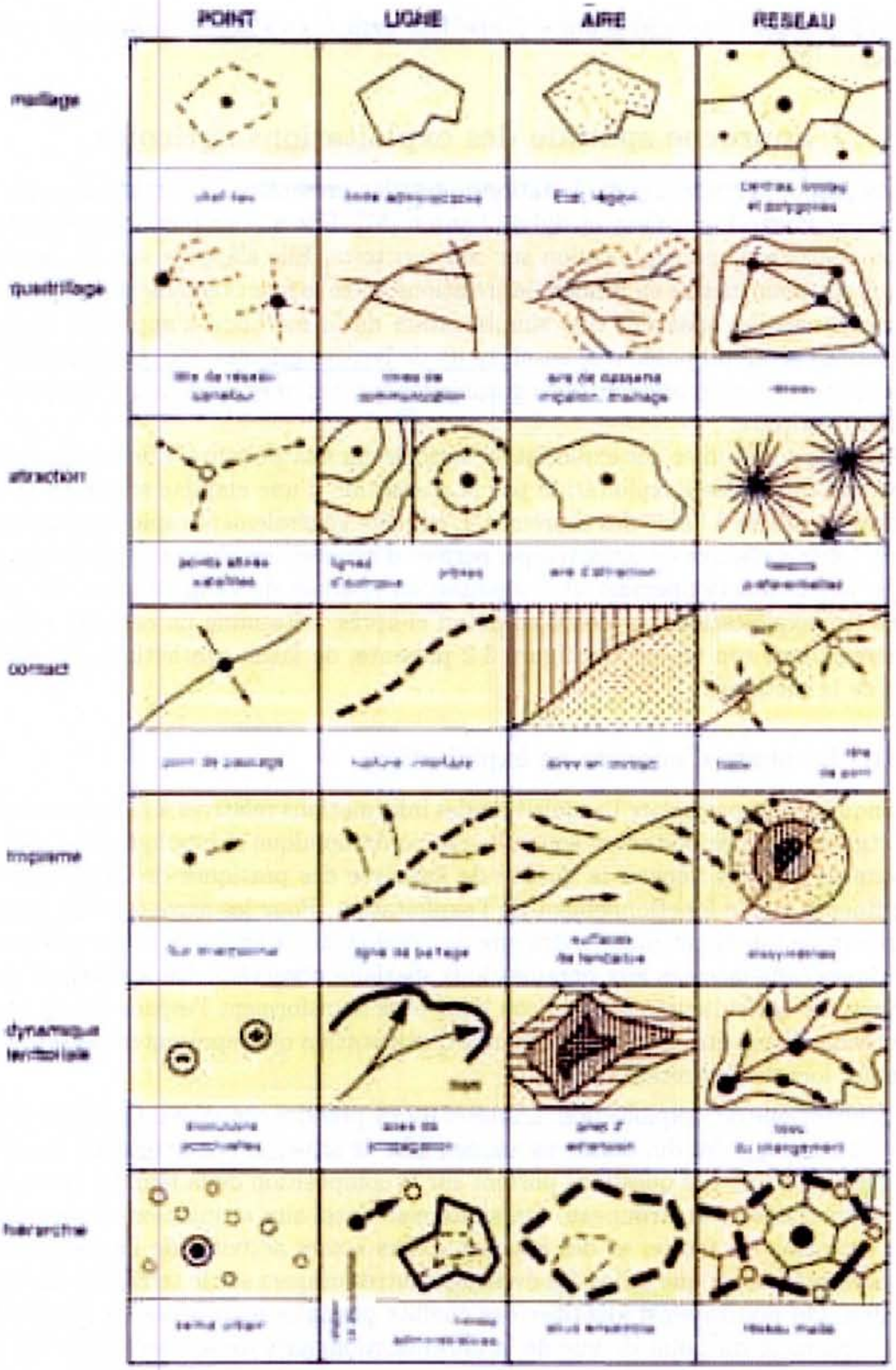


FIG. 3.1 – Tableau des 28 chorèmes élémentaires utilisés pour la modélisation graphique des régions [11].

duire la façon dont les sociétés ont organisé leur espace, en s'accommodant des contraintes physiques » [29].

3.2.3 L'approche spatiale des exploitations agricoles

La méthode d'étude des exploitations agricoles présentée a été mise au point et formalisée par Pierre-Louis Osty et Sylvie Lardon [57]. Elle a pour but d'appréhender l'inscription spatiale d'une exploitation sur son territoire. Elle s'appuie sur les pratiques de l'agriculteur pour mettre en lumière les relations entre le fonctionnement de l'exploitation et son organisation spatiale. Une simplification de la méthode d'approche spatiale des exploitations a été proposée aux enseignants de lycées agricoles [99]. Cette méthode permet la production de représentations graphiques synthétiques modélisant les exploitations dans leur territoire.

Elle permet d'étudier une exploitation agricole dans sa globalité. Elle est fondée sur une observation directe de l'exploitation par la réalisation d'une enquête suivie de la production d'une synthèse à l'aide des chorèmes. L'enquête généralement rapide (maximum deux heures), parfois réalisée en deux temps, permet d'acquérir les éléments pertinents concernant la vie, le fonctionnement et l'organisation spatiale de l'exploitation. Le protocole d'étude des exploitations, présenté en détail ci-après, détermine un canevas pour guider l'enquêteur dans son travail. La figure 3.2 présente, de façon synthétique, les différentes étapes de la méthode.

3.2.3.1 La phase d'enquête en exploitation

L'enquête doit permettre l'acquisition des informations relatives à l'exploitation et aux interactions entre l'exploitant et son milieu socio-économique et biophysique. De la qualité des données relevées dépend la qualité de l'analyse des pratiques de l'agriculteur et de leur influence sur le fonctionnement de l'exploitation. Pour les agronomes, il importe de comprendre le fonctionnement technique des exploitations agricoles et de rendre compte des logiques sous-jacentes aux organisations spatiales observées. On s'intéresse donc aux pratiques des agriculteurs et à la façon dont elles transforment l'espace ou s'y adaptent. Le protocole d'enquête définit cinq champs d'exploration qui représentent autant d'étapes à dérouler lors de l'entretien :

1. L'historique de l'exploitation. L'intérêt est de prendre connaissance de l'histoire, particulièrement des dix dernières années, afin de situer le contexte dans lequel évolue l'exploitation. Les questions portent sur la composition de la famille, l'évolution des productions et du troupeau. On s'intéresse aussi aux acquisitions de matériel et à l'évolution du foncier et des bâtiments. Les autres activités de l'agriculteur seront envisagées ainsi que celles des éventuels autres usagers sur le territoire de l'exploitation. On pourra ainsi identifier des conflits possibles mais aussi les implications de l'exploitant du point de vue de la multifonctionnalité de son territoire. Ce premier champ d'investigation permet de cerner l'exploitation dans son contexte naturel, social, économique, historique...
2. La structure du territoire. Ce deuxième champ d'investigation a pour objet de situer l'exploitation dans le territoire. L'enquêteur s'appuie sur les documents qu'il aura

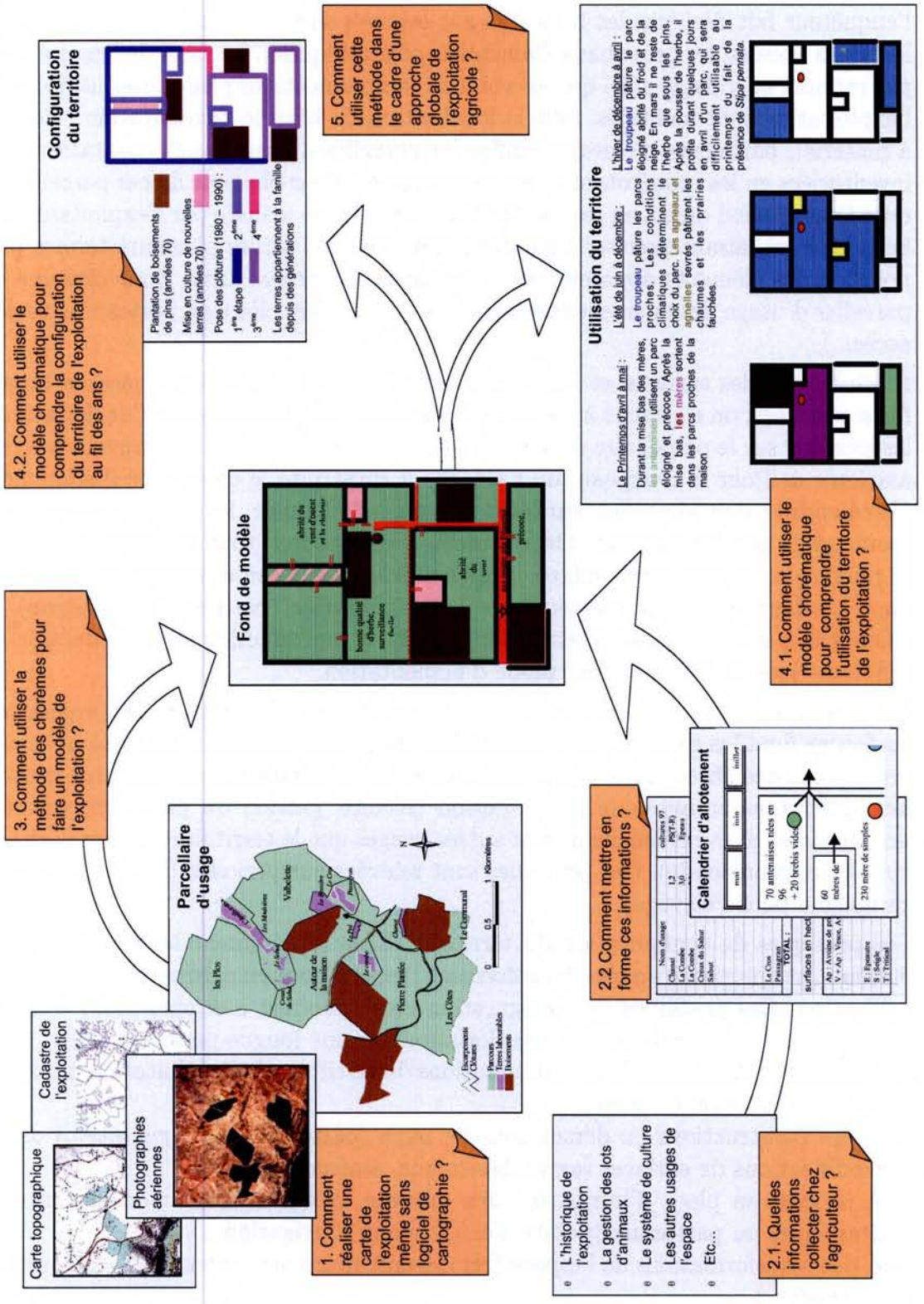


FIG. 3.2 – Plaquette présentant les étapes de la modélisation d'une exploitation agricole à l'aide de la méthode d'approche spatiale de l'exploitation agricole [99].

élaborés auparavant pour interviewer l'exploitant. Un fond de carte préparé avec le parcellaire de l'exploitation sert de base à la discussion. Dans un premier temps, l'enquêteur fait désigner les lieux-dits sur lesquels se situe l'exploitation et repérer les zones boisées, les principaux éléments topographiques... Le siège de l'exploitation par rapport au village, ainsi que les voies de communications principales utilisées par l'exploitation sont identifiés. Suit la localisation du bâti de l'exploitation : hangars à matériel, bâtiments d'élevage. Enfin, les parcelles d'usage de l'exploitation, sont inventoriées en les numérotant et en les nommant. Pour chacune de ces parcelles, on note l'occupation du sol en respectant la nomenclature utilisée par l'exploitant et au besoin en lui faisant préciser les correspondances, s'il utilise plusieurs termes pour nommer des éléments apparemment identiques. On relève aussi, pour chacune des parcelles d'usage, les aménagements (points d'eau, silos, clôtures, portes éventuelles, accès...).

3. L'assolement des surfaces et l'allotement des animaux. En ce qui concerne les surfaces cultivées, on s'attache à identifier les successions présentes sur l'exploitation, à les localiser sur le parcellaire d'usage et à comprendre la logique qui prévaut pour cet assolement. Pour le troupeau, on s'attache à construire le calendrier d'allotement. Ce calendrier doit identifier, sur les douze mois de l'année, les différents lots d'animaux gérés par l'exploitant. On s'attache à connaître à tout moment de l'année, le nombre et la catégorie (mères, jeunes de l'année...) des animaux de chaque lot ainsi que leur position sur le territoire de l'exploitation (pâturage, stabulation). Ce calendrier permet aussi de schématiser le cycle de reproduction et la production des différents lots en lien avec leur mode d'alimentation.
4. Les pratiques d'utilisation du territoire. Les pratiques d'utilisation de l'espace sont les façons dont les éleveurs, au cours d'une campagne, mobilisent les ressources pour répondre à des objectifs de production. Elles correspondent aux pratiques relatives aux différentes activités de l'exploitation (récolte, gestion du pâturage, entretien des surfaces) et éventuellement aux autres usages que le territoire supporte (chasse, cueillette, promenade). Ces pratiques sont relevées par périodes (e.g. saisons d'une campagne pour l'élevage).
5. Les pratiques de configuration du territoire. Elles représentent les façons dont les éleveurs, sur plusieurs années, transforment l'espace pour en permettre une meilleure utilisation. Ces pratiques de configuration correspondent aux opérations foncières et d'aménagement réalisées sur l'exploitation sur une longue période :
   - À des agrandissements ou des diminutions du territoire de l'exploitation par achat, vente ou location de terres,
   - À des constructions ou démolitions de bâtis (bâtiment d'élevage, gîte...) ou des modifications de surfaces (agrandissements, aménagements),
   - À la mise en place d'infrastructures (chemin, clôtures...), d'aménagements hydrauliques ou pastoraux (points d'eau, canaux d'irrigation...),
   - À des transformations de l'espace (remise en culture, arasement de haies, création de fossés...).

Cette phase d'étude des pratiques de l'exploitation peut faire l'objet d'un deuxième entretien d'enquête. Dans ce cas, le temps entre les deux rencontres est mis à profit

pour mettre au propre les données accumulées sur le parcellaire d'usage. Cette phase est l'occasion de corriger ou de préciser certaines perceptions du parcellaire recueillies lors de la phase précédente.

3.2.3.2 La phase de synthèse graphique

Au terme de l'enquête, l'agronome compile ses observations et traduit l'organisation spatiale de l'exploitation agricole sous la forme d'une représentation schématique synthétique. Le chorème d'exploitation est construit en trois étapes.

1. La structure du territoire. Après avoir réalisé l'enquête, les agronomes s'attachent à identifier les objets spatiaux qui semblent être structurants pour le territoire de l'exploitation : des routes, une rivière, des zones de bâti. A ces premiers éléments retenus, s'ajoutent ceux qui ont du sens par rapport au fonctionnement de l'exploitation : cela peut être de la forêt, des blocs de parcelles recevant des cultures annuelles, des pâtures ayant un rôle clé dans la conduite des animaux... Cette sélection et description de l'occupation du sol sur le territoire de l'exploitation agricole est importante. C'est l'étape où l'on passe de la carte au modèle. Le schéma n'est pas exhaustif. Seules les informations qui ont un sens pour l'organisation du territoire de l'exploitation sont reportées. La nomenclature des objets spatiaux retenus est adaptée en fonction des différentes natures d'occupation du sol rencontrées

2. L'utilisation du territoire. A partir des données recueillies dans l'enquête et des objets retenus après simplification, une correspondance entre la structure et l'utilisation est établie. La structure correspond aux objets retenus, à leur position relative et à leur signification. Leur utilisation explicite leur rôle dans le fonctionnement de l'exploitation. Par exemple, la pâture des vaches laitières du même côté de la route que le bâtiment d'élevage est une structure qui a pour fonction de faciliter la circulation des animaux. On met ainsi en relation les objets avec les règles de fonctionnement énoncées par l'agriculteur. Plus que le constat de l'existence d'une structure spatiale, c'est la reconnaissance des mécanismes sous-jacents que l'on recherche. Les utilisations sont reportées sur le fond de modèle de la structure, éventuellement en plusieurs périodes.

3. La configuration du territoire. Enfin, les éléments correspondant aux opérations foncières et aux aménagements réalisés sur l'exploitation sont reportés de façon simplifiée. Ces éléments correspondent aux pratiques de configuration du territoire sur une longue période. Ils peuvent être regroupés par périodes et reportés sur plusieurs copies du schéma de structure obtenu précédemment.

La méthode, présentée ici, est à adapter à chaque problématique et à chaque situation. Elle peut être utilisée pour étudier l'ensemble des exploitations d'un territoire. La section suivante présente l'utilisation de cette méthode lors d'une semaine de terrain qui s'est déroulée début avril 2001 sur le causse de Sauveterre au nord des gorges du Tarn et plus précisément au lieu-dit Soulages, sur la commune de St Georges de Lévéjac (figure 3.3).

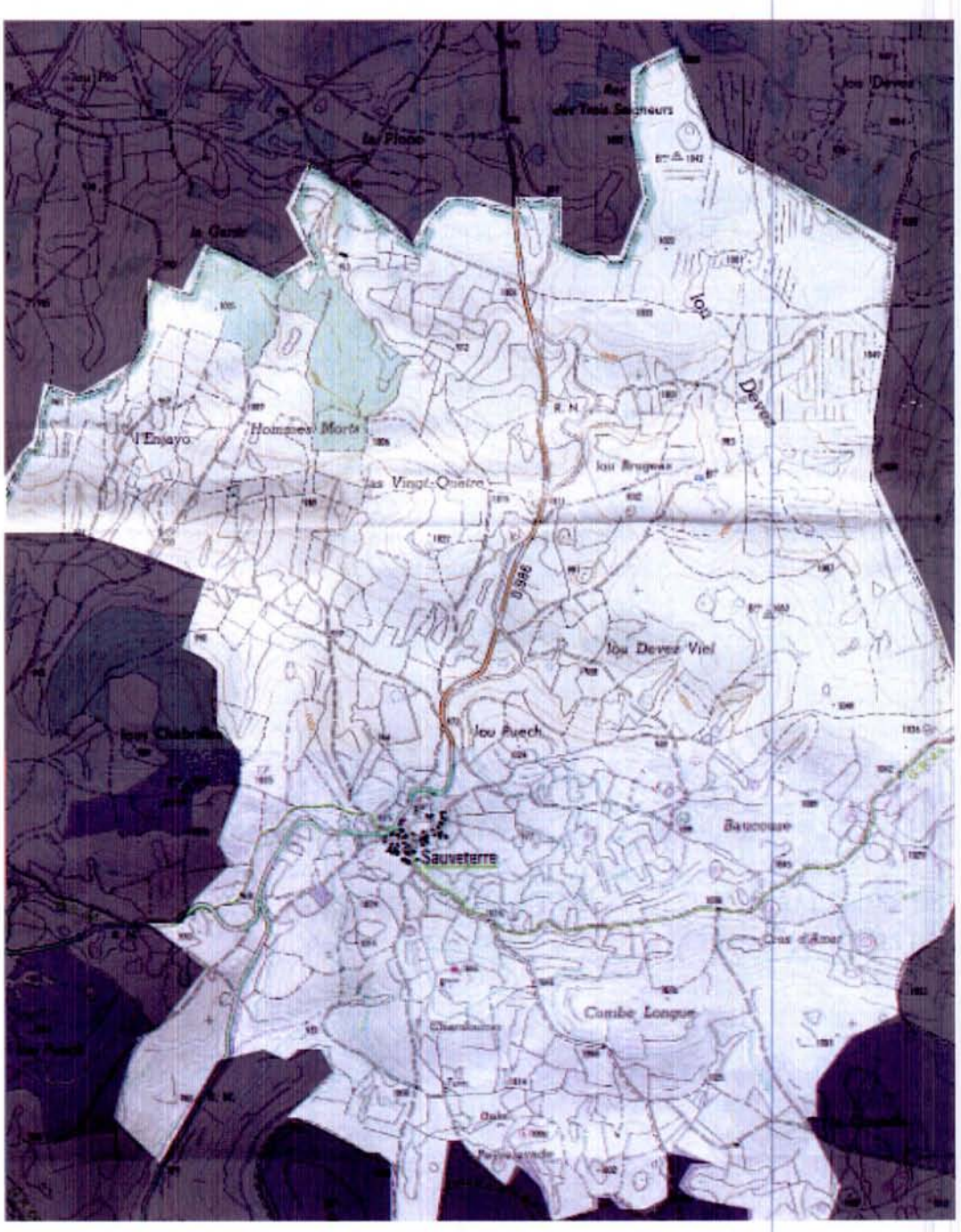


FIG. 3.3 – Carte du causse de Sauveterre

3.3 Exemple d'application de la méthodologie : le cas de l'étude du Sauveterre

L'expérience relatée ici s'inscrit dans un programme de recherche sur la dynamique du paysage au sud du Massif Central [77]. Il s'agit d'étendre les connaissances sur les exploitations agricoles du causse Méjan au causse de Sauveterre, plus au Nord, mais présentant les mêmes caractéristiques de milieu bio-physique, en particulier pour les risques de fermeture de paysage. En prolongement de la comparaison des résultats obtenus au niveau des exploitations agricoles, les agronomes s'intéressent aussi à l'organisation des lieux-dits, vus comme la combinaison des territoires d'exploitations, sur l'ensemble du causse.

Un échantillon représentatif des différentes situations observables sur le causse de Sauveterre a permis de définir la zone d'étude. Elle couvre une bande de territoire constituée de 20 lieux-dits contigus, allant du causse boisé au causse nu, suivant un axe ouest-est. Toutes les exploitations ont été enquêtées, sauf celles qui ont refusé l'enquête, correspondant à un effectif total de 24 exploitations (sur les 30 présentes).

Une étape préalable a consisté à recueillir des informations générales sur l'ensemble du causse, des lieux-dits et des exploitations à enquêter. La méthodologie d'enquête a été simplifiée et testée au préalable ; la structure des lieux-dits a été repérée par un passage sur le terrain. L'enquête relative à une exploitation laitière du lieu-dit Soulages est rapportée et le chorème d'exploitation produit est présenté. Cela permet à l'informaticien une première prise de contact avec les questions des agronomes, qui sont ensuite récapitulées.

3.3.1 Étape préparatoire à l'enquête

La réalisation d'une enquête efficace passe par une étape préparatoire. Cette étape a pour but de collecter des données sur la région où se situe l'exploitation visée, mais aussi des données propres à l'exploitation. Préalablement à l'enquête en exploitation, un travail de collecte et d'analyse de données générales est effectué. Des données statistiques fournies par les services de l'État, le SCEES[17] et la DDAF sont consultées pour la zone sélectionnée. Elles permettent d'avoir une idée de la répartition des exploitations sur la région, de leur nombre et des propriétaires des terres. Enfin, pour chaque exploitation à enquêter (lorsque les exploitants sont présents, disponibles et volontaires pour ce genre d'exercice), un recueil d'informations préalables est effectué à partir des parcellaires des déclarations PAC[18] et mis en relation avec les données issues de la lecture des cartes IGN[19] au 25 millième et des photographies aériennes disponibles.

Ainsi, le travail de préparation a permis de réunir les informations nécessaires sur les différentes exploitations. Elles ont été reportées sur des fonds de carte IGN pour chaque exploitation présente sur le territoire étudié.

Quatre personnes constituent l'équipe d'enquêteurs : M. Capitaine, S. Lardon, M. Naïthlo et P.-L. Osty. Trois d'entre eux connaissent bien la région et ont déjà travaillé

[17]Service Central des Enquêtes et Études Statistiques du Ministère de l'Agriculture et de la Pêche.

[18]Pour bénéficier des aides européennes à l'agriculture, les exploitants doivent chaque année faire une déclaration de leurs surfaces exploitées en précisant leurs utilisations.

[19]Institut Géographique National.

sur des études menées sur le causse Méjan, le quatrième (M. Capitaine) a par contre travaillé en Lorraine, mais avec le même type d'approche [16]. Cette bonne connaissance générale du terrain leur permet d'envisager la réalisation d'enquêtes rapides, centrées sur la compréhension des grandes structures de chaque exploitation. Les détails sur les exploitations sont trop lourds pour le travail envisagé qui s'intéresse à l'activité d'un nombre important d'exploitations sur un même territoire. Aussi, pour adapter la méthode d'étude des exploitations à ce contexte, une enquête-test a été menée la semaine précédant notre intervention. Cette enquête a permis de mettre au point le questionnaire, d'organiser la prise de notes et de définir le déroulement des entretiens. L'exposé des résultats de cette enquête a aussi pour but de mettre d'accord les enquêteurs, qui travailleront en binôme, afin d'avoir une homogénéité des informations recueillies. Ainsi, quatre phases ont été définies pour diriger les entretiens en précisant le type d'informations à recueillir ainsi que l'enchaînement souhaité des questions. Ces phases se développent comme suit :

1. La structure du territoire de l'exploitation est décrite par l'exploitant. Le parcellaire reproduit préalablement sur un fond de carte sert de soutien au discours de l'exploitant.
2. Le calendrier d'allotement est alors collecté et l'utilisation du territoire est précisée. Ce deuxième point de vue sur le territoire de l'exploitation permet à l'enquêteur d'organiser et de compléter les informations acquises précédemment. Cette phase permet d'avoir une vision de l'utilisation du territoire au cours de l'année.
3. Suit le questionnement sur l'historique de l'exploitation et de la configuration de son territoire. La période d'installation est souvent une bonne entrée pour cette phase, mais c'est l'évolution au cours des dix dernières années qui est surtout prise en compte. L'entretien est dirigé afin d'appréhender les techniques de valorisation et les stratégies d'investissement développées (gestion des cailloux, défrichage, etc.). Cette phase est encore mise à profit pour compléter les informations recueillies lors de la phase 2. Elle permet de comprendre les pratiques de configuration du territoire qui ont l'impact le plus important sur le paysage.
4. Enfin les dernières questions, en guise de conclusion, cherchent à déterminer les enjeux sur le territoire du lieu-dit, et la façon dont sont abordés les problèmes liés à l'environnement et au multi-usage de l'espace par l'exploitant.

Ces enquêtes durent entre une et deux heures en fonction des exploitants, mais en règle générale, un entretien dure autour d'une heure et demie. Les enquêteurs, par équipe de deux se répartissent la prise de note et le questionnement, de façon à pouvoir faire l'enquête en une seule fois. L'un vérifie que les informations recueillies sont complètes sur un thème pendant que l'autre enchaîne avec les questions liées à la phase suivante. De plus, les enquêteurs ont toujours à l'esprit de recouper le plus possible les informations recueillies afin d'en assurer la cohérence.

3.3.2 L'enquête en exploitation

Après avoir pris connaissance du protocole d'enquête, nous nous rendons sur le lieu-dit Soulages (figure 3.4). Une lecture de paysage est effectuée par les membres de l'équipe

pour avoir une bonne connaissance du secteur. Nous repérons plusieurs zones, lieux de différentes activités agricoles, autour du village, qui nous serons expliquées ensuite par les agriculteurs enquêtés.

Tout d'abord, il y a une zone dénommée la "Lande" située à l'ouest du village. C'est un bas-fond où l'eau est présente toute l'année. Un fossé la sépare en deux parties. Les bergers y descendaient leurs bêtes lorsqu'ils les emmenaient sur les sectionnaux proches. Cette zone humide est qualifiée, par la plupart des exploitants, comme possédant les meilleures terres du village (voire du causse de Sauveterre). De nombreux propriétaires se la partagent, sous forme de petites parcelles étroites perpendiculaires au fossé. Cette zone fait l'objet d'un usage particulier, institué vers 1760 par le Seigneur de Séverac et toujours en vigueur : la première coupe d'herbe revient au propriétaire ou à son fermier ; la seconde est réservée, sous le régime de la veine pâture, aux "bêtes à cornes" ; la troisième est destinée, sous le même régime et à partir du 29 septembre, aux vieilles brebis. Au nord du village, se situe la zone de sectionnaux de la Devèse. Le même type de régime d'utilisation que pour la Lande y a persisté jusqu'à la sécheresse de 1976. Au sud-est du village, se dessine la "Combe" constituée principalement de terres labourables réparties entre les différents exploitants de la commune.

Cette lecture permet de donner quelques points de repères utiles pour comprendre rapidement les exploitants lors des entretiens. Elle permet de se faire une première idée de la structure du lieu-dit, qui servira de fond pour positionner les exploitations enquêtées.

Nous présentons ici le compte-rendu d'enquête d'une exploitation en GAEC[20] associant trois exploitants : le fils, son père et sa mère. Ils sont aidés dans leur travail par un frère et un aide familial.

Une des premières remarques de l'exploitant pose le problème ainsi : "Pourquoi des exploitations agricoles sur le causse, là où personne ne voudrait des terres ?" La réponse vient dans la foulée "La société de Roquefort". En effet, grâce au Roquefort, le lait de brebis est fortement valorisé et permet le développement des exploitations en place.

L'enquête commence par la structure du territoire de l'exploitation. L'exploitant déclare travailler sur 175 hectares mais ne nous décrit l'utilisation que de 160 ha. Ses terres se répartissent sous forme de sectionnaux pour 40 ha en 4 parcelles et 8 ha de bois et environ 110 ha découpés en parcelles de 1 ha en général, la plus grande faisant 5 ha. L'exploitation est propriétaire de 45 ha, 35 ha sont en fermage au nom du GAEC alors que le reste est en fermage au nom du père.

La répartition des surfaces de l'exploitation se décompose comme suit :

- 22 ha sont consacrés à la culture de céréales. Ce sont des terres labourées principalement dans la "Combe" et dans la "Lande". Le morcellement y est important car chaque exploitant en possède un peu ;
- 50 ha sont consacrés au fourrage. Ce sont des prairies semées et renouvelées tous les 4 ou 5 ans en rotation avec les céréales ;
- 30 ha de prairie sont cultivées hors assolement. Elles sont renouvelées tous les 8 à 10 ans ;

[20] Groupement Agricole d'Exploitation en Commun. Le GAEC est une société civile de personnes permettant, à des agriculteurs associés, la réalisation d'un travail en commun dans des conditions comparables à celles existant dans les exploitations de caractère familial.

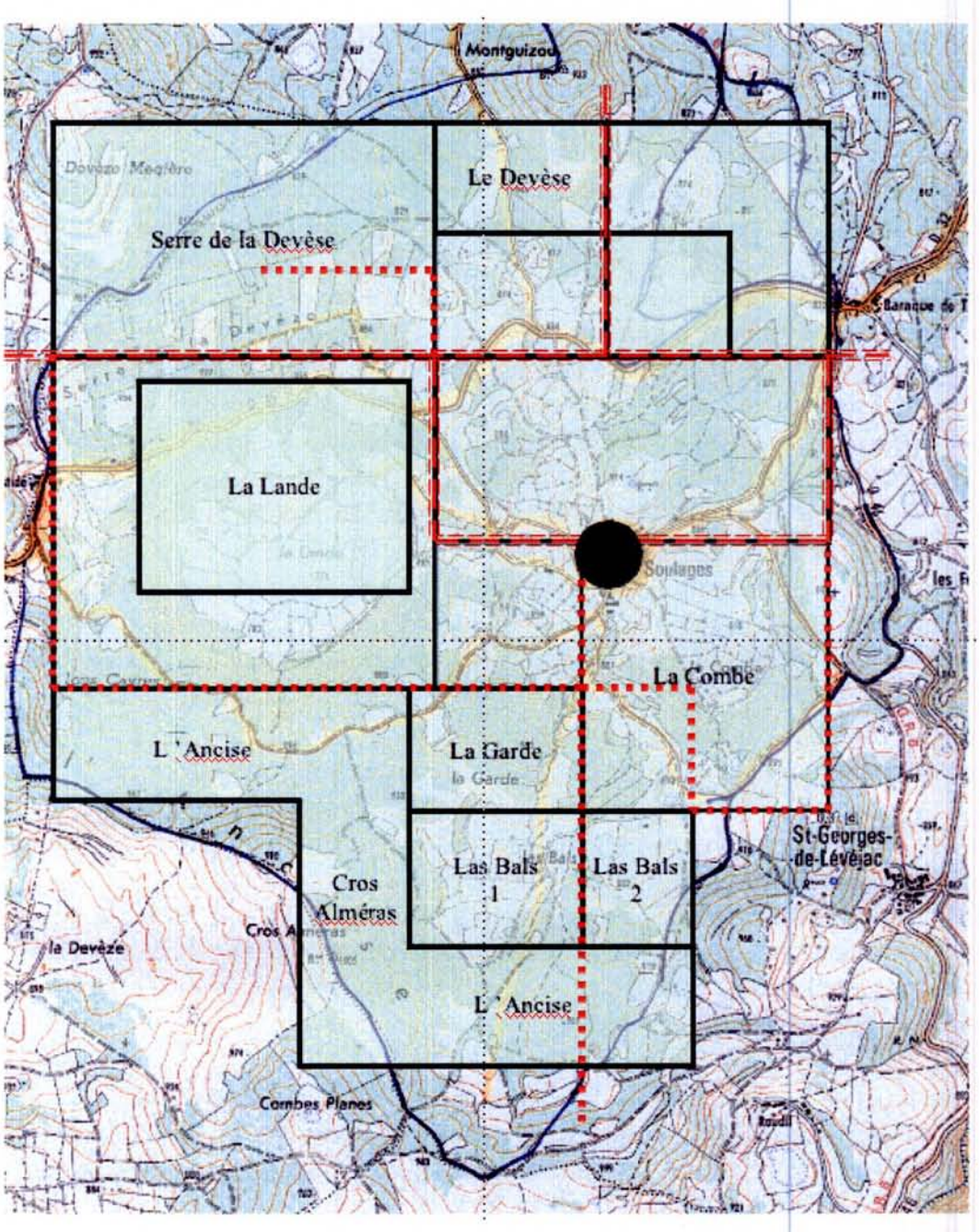


FIG. 3.4 – Territoire du lieu-dit de Soulages. Extrait carte 1 :25000 série bleu 2635O de l'IGN.

- 60 ha sont des parcours voire de bonnes pâtures. 75 à 80% de ces surfaces sont fertilisées à faible dose.
- 8 ha environ de son territoire ne sont jamais utilisés.

Les sectionnaux sont normalement des surfaces collectives. Des discussions, amorcées il y a une dizaines d'années, entre les différents exploitants utilisateurs de l'espace ont abouti, il y a cinq ans, à leur découpage en lots. Cette séparation en lots "privés" et l'obtention de subventions ont permis une mise en valeur de ces surfaces laissées à l'abandon et destinées à la friche. Les 4 parcelles qui ont été attribuées au GAEC ont été séparées par des haies et clôturées. Les clôtures sont disposées dans les haies. Des subventions ont permis un travail de nettoyage des sous-bois et une amélioration des parcours (défrichage). Les conséquences de ce travail d'aménagement sont que les parcours sont maintenant d'un entretien facile et qu'il sont amendables[21] sur leur presque totalité.

La redistribution des sectionnaux ainsi que les travaux d'aménagement des parcelles a aussi pour conséquence la disparition d'un grand nombre de chemins qui ont été intégrés dans les surfaces des exploitations à l'occasion des réaménagements effectués.

Une fois l'inventaire des surfaces de l'exploitation effectué, c'est leur utilisation qui est étudiée, notamment à travers la conduite du troupeau. La seule production de la ferme est le lait à raison de 1000 hectolitres de quota. L'ensemble des céréales, soit environ 80 tonnes, et du foin produits sur l'exploitation sont consommés. Un complément alimentaire est acheté à raison de 11 tonnes sur l'année.

Au mois d'avril le troupeau est constitué de 480 brebis adultes, 109 agnelles de renouvellement et 10 béliers. La période de lactation s'étale du 10 janvier à fin juillet (voire août certaines années). Début juillet, 400 brebis sont luttées alors que la lutte des agnelles se fait début août. Les mises bas ont lieu en décembre pour les adultes et en janvier/février pour les agnelles. Les agneaux sont vendus à 12 kg, les brebis de réforme sont sorties du troupeau vers le 11 mai (avant la lutte).

Au printemps, jusque fin avril, le troupeau est sorti, à proximité de la bergerie une quinzaine de jours pour que les jeunes bêtes apprennent à pâturer avec leurs aînées. C'est le père qui est le berger de l'exploitation.

Ensuite, le troupeau est séparé en trois lots : les agnelles de l'année, les antenaises (agnelles de l'année précédente mises à la reproduction) et les brebis mères. Les agnelles sont placées à 2 kilomètres de l'exploitation de juin à novembre dans un abri où elles sont rentrées tous les soirs. Elles pâturent à proximité, sauf les jours de grosse chaleur où elles pâturent la nuit. Cette conduite de troupeau est liée à la peur des chiens errants durant la nuit. Les autres bêtes (mères et antenaises) sont installées à la bergerie de l'exploitation. Elle sont conduites journellement sur les différentes parcelles de parcours autour de l'exploitation, le principal souci étant la taille des parcelles. En effet les parcelles de 1 ha, nombreuses, sont trop petites sur une journée. Les parcelles de 2 ha sont correctes, mais l'idéal serait d'avoir des parcelles d'environ 3 ha.

Pour fonctionner, cette exploitation est équipée de 4 tracteurs, d'un broyeur de cailloux, d'un girobroyeur et d'une presse à grosses balles pour conditionner le foin. Les pratiques d'ensilage entre 1978 et 1981, puis d'enrubannage du foin encore humide entre 1982 et

[21] Les amendements sont des épandage d'engrais, de fumier sur des surfaces agricoles accessibles aux tracteurs.

1993 ont été abandonnées au profit du foin pour des soucis de qualité de la production de lait. Les aménagements fonciers n'ont pas été étudiés.

En terme de conclusion, le but de l'exploitant est de baisser autant que possible le coût de revient du lait. Pour atteindre cet objectif, la pratique de pâturage extensif du troupeau est mise en avant. Ceci implique que le troupeau possède une taille maximale au regard du quota de production et de la consommation d'aliments achetés en complément de l'auto-production.

3.3.3 Le chorème d'exploitation

Le soir même après les enquêtes, les notes de la journées sont organisées et une première ébauche des chorèmes des exploitations visitées est dessinée. Pour élaborer le chorème d'une exploitation, les experts confrontent leurs points de vue et compare les exploitations entre elles.

Le chorème de l'exploitation précédemment décrite est présenté figure 3.5. Il montre une exploitation avec une utilisation relativement intensive de l'espace, liée à la forte proportion des terres assolées et de parcours fertilisés, l'existence de parcs pour la mise à l'herbe et la combinaison de champs et de parcs à proximité de la bergerie principale. Les surfaces éloignées, au sud du lieu-dit, sont gérées à distance pour les animaux à faibles besoins (ici les agnelles à partir de l'été), grâce à l'existence d'un abri et à l'allotissement des sectionnaux qui a fortement contribué à la configuration de cet ensemble à distance.

Les riches discussions lors des soirées durant la semaine de terrain et celles qui ont suivi pour synthétiser le travail des ces enquêtes prouvent la nécessité d'ajustements entre les enquêteurs pour trouver les formulations graphiques qui conviennent. D'autant que la vision de ce schéma, n'en permet pas la lecture aisée pour une personne extérieure. Enfin, les interprétations possibles du fonctionnement de l'exploitation, à la lecture du chorème, ne sont possibles que pour des experts ayant une bonne connaissance du terrain et une expérience commune avec les modélisateurs ayant produit le chorème.

Ces remarques nous poussent à poser quelques questions sur cette méthodologie de l'étude de l'espace. Nous explorons ces questions et quelques autres au chapitre suivant.

3.4 Questions soulevées par la méthodologie appliquée

Les chorèmes permettent d'extraire les principes organisateurs de l'utilisation de l'espace par les exploitations. Ils donnent à voir de façon synthétique les modalités d'utilisation du territoire. Cependant, les chorèmes produits dépendent fortement des experts qui les conçoivent et des situations étudiées. Il reste une grande part d'ambiguïté dans l'expression même du graphique. La grille des chorèmes élémentaires donne une procédure de construction, mais elle ne permet pas d'évaluer les variations observées. Les modes d'usage de cet outil ne sont pas encore totalement fixés.

Ces remarques poussent à se poser la question de la reproductibilité de cette méthode de modélisation spatiale des exploitations agricoles fondée sur l'utilisation des chorèmes. En effet, l'on a pu constater que les chorèmes produits dans le cadre des travaux sur le

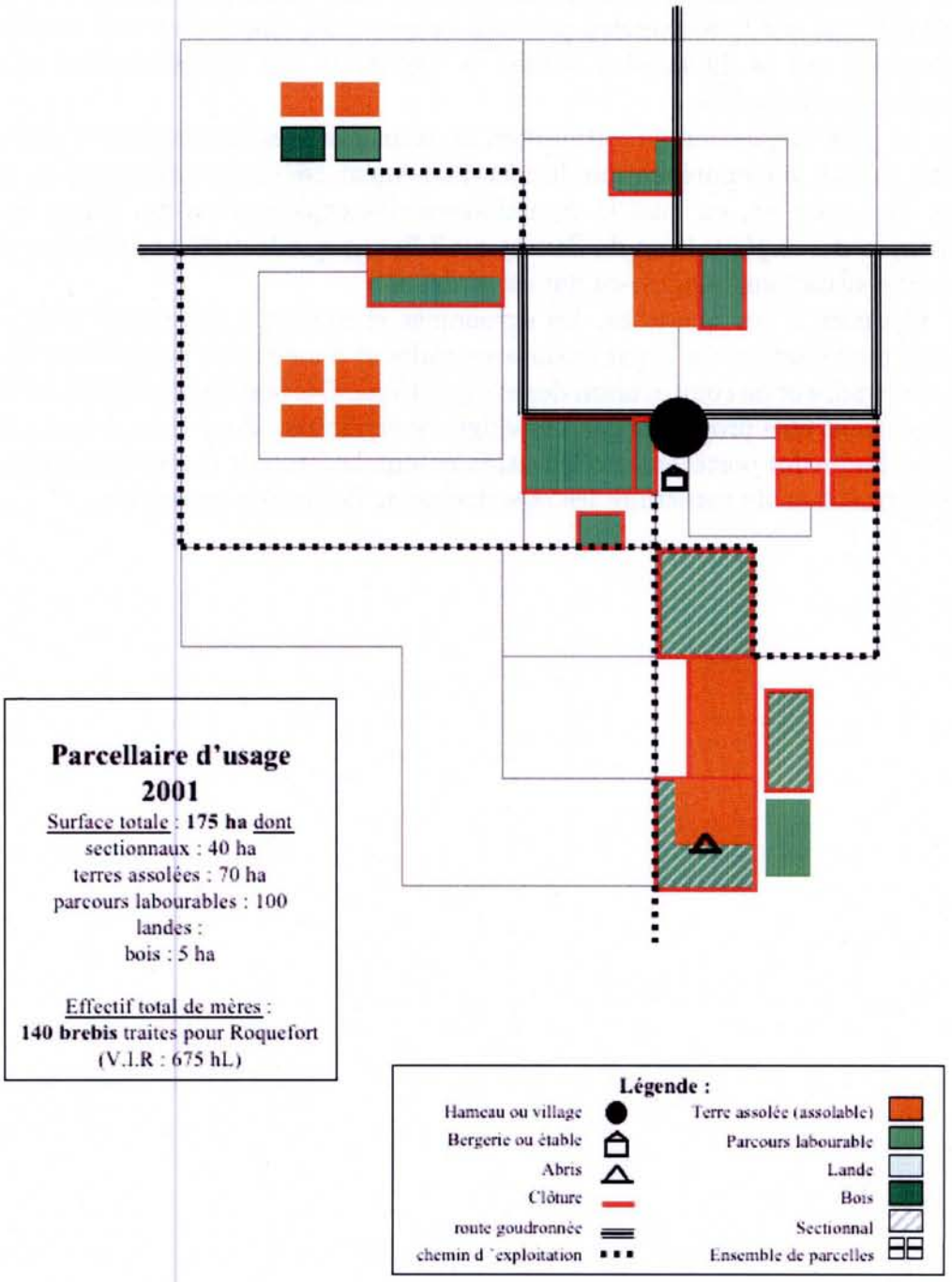


FIG. 3.5 – Exemple de synthèse sous forme de chorème d'une exploitation enquêtée sur le causse de Sauveterre

périmètre de Vittel et sur les exploitations du causse Méjan sont assez différents dans leur représentation. Pourtant, les fondements sont les mêmes.

Même lorsque la méthode et la problématique sont identiques, comme dans le cas du travail effectué sur le Sauveterre présenté ci-avant, les discussions sont encore riches d'enseignements sur la difficulté à réduire la variabilité des représentations et des interprétations fonctionnelles.

Enfin, se pose la question de l'utilisation de connaissances acquises sur d'autres territoires pour aider à la compréhension du fonctionnement des exploitations sur un nouveau territoire. Par exemple, en quoi la connaissance des exploitations du Méjan facilite la compréhension des exploitations du Sauveterre ? Est-ce que la méthode est reproductible dans d'autres situations ? Qu'est-ce qui est stabilisé ?

Pour répondre à ces questions, les agronomes cherchent à formaliser plus avant la méthode de modélisation des organisations spatiales et à mettre au point des outils de validation, de mesure et de comparaison des exploitations. Ces besoins font appel à l'étude de solutions pouvant être proposées par l'intelligence artificielle. Pour cela, il faut définir les données sur lesquelles porteront modélisation et simulation puis capter les connaissances associées permettant de construire les raisonnements devant s'y rapporter.

Chapitre 4

Modèle de RàPC pour les organisations spatiales agricoles

Sommaire

4.1 Le processus d'acquisition 68
4.2 Application du RàPC à la problématique agronomique . 70
4.3 Des graphes pour modéliser les organisations spatiales . . 71
4.4 Un modèle des connaissances spatiales 72
4.4.1 Les régions 73
4.4.2 Les relations 74
4.4.3 Région ou Relation ? 77
4.5 Autres connaissances du domaine 79
4.5.1 Les explications 79
4.5.2 Les points de vue 80
4.5.3 Les échelles 81
4.6 Conclusion 81

Dans le chapitre précédent, nous avons vu comment les agronomes envisagent l'étude des exploitations agricoles et leur inscription spatiale dans leur environnement à travers une démarche d'analyse globale. Nous avons décrit comment ils synthétisent leurs connaissances sur les exploitations étudiées dans des représentations graphiques : les chorèmes.

Dans ce chapitre, nous développons notre proposition pour la construction d'un système à base de connaissances, fondé sur le raisonnement à partir de cas, pour répondre à leurs attentes. Puis, nous expliquons comment nous traduisons les organisations spatiales et le fonctionnement des exploitations dans le système envisagé. Nous détaillons le modèle sous-jacent utilisé pour représenter les connaissances sur les exploitations. Enfin, nous présentons comment s'est construit la base de connaissances du domaine en interaction avec les agronomes.

4.1 Le processus d'acquisition

Au sens informatique, acquérir de la connaissance consiste à formaliser la sémantique dans un format syntaxique qui pourra dès lors être manipulé de façon calculatoire par des ordinateurs.

Les connaissances concernant un problème donné constituent un modèle de ce problème et de son environnement. Selon le type de connaissances mobilisées par le modélisateur, on distingue deux types de modélisations [3] :

- Informelle. Elle s'effectue en terme de connaissances phénoménologiques. Elle se manipule en fonction du sens des termes et non en fonction de leur forme syntaxique. Le rapport entre modèle et réalité correspond au signifié du modèle.
- Formelle. Elle s'effectue en terme de connaissances scientifiques offrant des capacités prédictives. Deux formes de modélisation formelle existent :
  - La mathématisation se fonde sur la géométrie et le continu. Le modèle construit mathématiquement la réalité ;
  - La formalisation se fonde sur le calcul symbolique. Le modèle dénote la réalité.

Ainsi les données et connaissances manipulées par les agronomes doivent être formalisées pour permettre leur représentation au sein du futur système, afin de les utiliser dans les mécanismes du raisonnement à partir de cas. Ces données et connaissances sont de différents types : données d'enquêtes, chiffres, cartes, synthèses écrites et graphiques, expériences du terrain, expertises sur les régions et exploitations étudiées.

Le travail d'acquisition associe agronomes et informaticiens. Il s'appuie sur la transformation des chorèmes en graphes. Les chorèmes d'exploitation sont choisis comme support parce qu'ils sont déjà des synthèses des connaissances sur des exploitations. Il sont de plus le résultat du raisonnement des agronomes qui relient l'organisation spatiale observée à des modalités de fonctionnement des exploitations.

Les systèmes de reconnaissance des formes sont de plus en plus performants. Il existe de nombreux systèmes tant pour la lecture de plans d'architectes, ou d'images plus complexes encore [84], comme les cartes routières. Un système de reconnaissance de structures de paysage a aussi été développé dans le cadre des travaux précédents menés en partenariat entre l'équipe Orpailleur et les agronomes du SAD de Mirecourt [79, 80]. Aussi, l'idée de faire de la reconnaissance automatique de structures à partir des schémas chorématiques

a-t-elle été évoquée.

Mais notre problème ne peut être résolu de la sorte. En effet, nous nous intéressons au rapport entre la structure spatiale des exploitations et leur fonctionnement. Or, cette relation n'est en rien évidente à la lecture des compositions chorématiques. Ces schémas sont déjà une modélisation par l'agronome producteur du dessin de l'existant qu'il a observé. Cette modélisation est faite en fonction des éléments considérés comme importants pour comprendre le fonctionnement d'une exploitation donnée en rapport à son organisation spatiale. Le dessinateur fait le choix de représenter ou non certains éléments. De plus, les distances ne sont pas conservées. Le chorème n'est pas une carte. La position relative des éléments du chorème répond à des soucis de représenter des relations plus ou moins fonctionnelles. C'est en fait par le discours des agronomes que nous pouvons obtenir les informations nécessaires pour définir les relations entre les différentes entités du dessin des chorèmes. Finalement, notre objet n'est pas de faire de la reconnaissance automatique d'images sur les chorèmes, mais bien de "faire parler" les agronomes afin d'acquérir les connaissances nécessaires au système.

L'acquisition des connaissances et leur modélisation ont été menées au cours de plusieurs séances de travail réunissant agronomes et informaticiens. L'acquisition s'est ainsi faite par itérations successives de phases d'interactions entre agronomes et informaticiens, suivies de phases de maturation, offrant un temps pour la formalisation des connaissances acquises. Cette maturation est souvent l'occasion du développement de nouvelles pistes, de nouvelles solutions dans la pratique et l'utilisation des outils tant graphiques (chorèmes et graphes) que formels : construction d'un référentiel commun admis par tous pour décrire les connaissances du domaine.

En particulier, deux séances de confrontation agronomes-agronomes et agronomes-informaticiens ont été filmées dans le cadre d'une collaboration avec des psycho-cogniticiens et des linguistes. La première séance a notamment porté sur la formalisation des cas, i.e. la formalisation des problèmes et de leurs solutions. Pour la deuxième séance, nous nous sommes focalisés sur l'acquisition et la formalisation des connaissances d'adaptation.

La confrontation entre agronomes et informaticiens est un moyen de leur faire exprimer comment ils décrivent les structures spatiales qu'ils représentent dans leurs compositions chorématiques, mais aussi comment ils interprètent ces structures. C'est l'occasion de réinterroger leurs connaissances et en particulier de leur faire expliciter les connaissances implicites qu'il mobilisent lors des interprétations.

La première séance a eu pour but de vérifier la reproductibilité de la méthode de modélisation à l'aide des chorèmes en comparant les démarches appliquées par les agronomes sur des exploitations de la région de Vittel et sur des exploitations du causse Méjan. Les questions posées à cette séance étaient : « Est-ce qu'un agronome du Méjan est capable d'interpréter un chorème de Vittel et inversement ? » et « Est-il possible de traduire le chorème sous forme de graphe ? »

La deuxième séance de travail a eu pour but de déterminer les connaissances d'adaptation mobilisées dans le cadre d'un travail sur deux régions proches en terme d'environnement naturel. Elle a porté sur l'étude des systèmes d'exploitation du causse de Sauveterre en se fondant sur les connaissances importantes acquises sur les exploitations du causse Méjan. La question posée était : « En quoi les connaissances du fonctionnement d'une exploitation en regard de son organisation spatiale peuvent être utilisées pour expliquer

le fonctionnement d'une autre exploitation dont on possède seulement une description de la structure spatiale ? ». En d'autres termes : « Est-il possible de comparer des chorèmes et la connaissance de l'un sur l'autre ? » et « Peut-on aussi le faire avec des graphes ? »

Ces deux séances de collaboration filmées donnent lieu à une observation poussée et à une analyse encore en cours de développement de la part des psycho-cogniticiens et des linguistes [62, 63]. Il y a eu évidemment de nombreuses autres séances de travail entre agronomes et informaticiens qui ont porté sur la modélisation des connaissances (cas, concept, adaptation) et le test du système (chapitre 7).

4.2 Application du RàPC à la problématique agronomique

Les agronomes disposent de riches connaissances sur un certain nombre d'exploitations. Ils en ont produit des chorèmes d'exploitation décrivant leurs organisations spatiales. De plus, il sont en mesure d'interpréter ces chorèmes et de nous fournir des explications du fonctionnement des exploitations en rapport avec leur organisation spatiale.

Pour expliquer le fonctionnement d'une nouvelle exploitation, ils la décrivent en terme d'organisation spatiale, dans un nouveau chorème. Puis, à l'aide de leurs connaissances de terrain, de l'environnement de l'exploitation étudiée et des connaissances sur les exploitations précédemment étudiées, ils peuvent donner des explications de fonctionnement de cette dernière.

Ce mode de raisonnement est typique d'un *raisonnement à partir de cas*. Un *cas* est un problème accompagné de sa solution. Le *problème* correspond à ce qui est observable sur le terrain, l'agencement relatif des différents éléments structurant le territoire de l'exploitation. La *solution* correspond à ce qui peut être interprété, ce qui peut être compris du fonctionnement d'une exploitation ou d'une partie d'une exploitation agricole i.e. une explication. Dans le cadre de notre travail, un *cas* est défini ainsi :

Définition 1 (Cas) *Un* cas *est une unité d'expérience constituée d'une partie problème (une organisation spatiale) et de sa solution (une explication de fonctionnement).*

Nous pouvons préciser les définitions des *problème* et *solution* d'un cas de la façon suivante :

Définition 2 (Problème) *Un* problème *est représenté par une organisation spatiale particulière dont on veut trouver l'explication de fonctionnement associée.*

Définition 3 (Solution) *Une* solution *est une explication de fonctionnement associée à un problème défini par une organisation spatiale.*

L'ensemble des exploitations étudiées par les agronomes fournissent les premiers cas de la base de cas qui sera enregistrée dans le système à construire. Conformément à la méthodologie appliquée par les agronomes, le système doit, pour expliquer un problème donné, retrouver dans la base de cas un ou des cas similaires. Le problème à expliquer

correspond à l'*organisation cible* (noté `Org(cible)`). Le cas retrouvé est le *cas source* constitué de l'*organisation source* (noté `Org(src)`), similaire à l'organisation cible, et de l'*explication de source* (noté `Exp(src)`).

L'explication associée au cas trouvé `Exp(src)` sert alors de base pour proposer une explication au problème posé : `Exp(cible)`. Le premier mécanisme de recherche dans la base de cas correspond à la *remémoration* du modèle du RÀPC. Le deuxième mécanisme correspond à la phase d'*adaptation* du RÀPC.

Les opérations de remémoration et d'adaptation ne sont pas indépendantes. En effet, nous cherchons à retrouver des cas similaires qui soient adaptables afin de proposer une nouvelle explication. Pour cela, nous nous situons dans le cadre d'une remémoration guidée par l'adaptation. L'écart entre `Org(src)` et `Org(cible)` est représenté par un *chemin de similarité* [72]. Ce chemin est utilisé pour construire un *chemin de transformation* de la solution de source `Exp(src)` afin d'obtenir l'explication `Exp(cible)` de `Org(cible)`.

Les chemins de similarité comme les chemins de transformation sont calculés en s'appuyant sur les connaissances du domaine qui seront précisées plus loin dans ce chapitre.

Pour permettre l'application du RÀPC, il nous faut acquérir différents types de connaissances. Tout d'abord, il y des connaissances particulières sur les premiers cas à enregistrer dans la base de cas. Ces connaissances sont les descriptions des organisations spatiales des exploitations et la description des explications associées. Ensuite, il faut des connaissances générales, un modèle du domaine, pour permettre la remémoration grâce à des mesures de similarité. Cette connaissance est constituée de concepts organisés dans une hiérarchie. Enfin, des connaissances d'adaptation sont nécessaires afin d'envisager le cycle complet du RÀPC.

4.3 Des graphes pour modéliser les organisations spatiales

Nous proposons de traduire les organisations spatiales sous forme de graphes étiquetés. En effet, les graphes sont des représentations graphiques qui peuvent être comprises par les agronomes tout en offrant un formalisme implantable pour lequel il existe des outils informatiques pour leur manipulation. Notamment, nous avons à notre disposition des algorithmes permettant de calculer des similarités entre graphes à l'aide d'opérations élémentaires telles que l'ajout ou la suppression d'un sommet ou d'un arc [12, 87].

D'autre part, les organisations spatiales sont naturellement représentables par des graphes. Nous avons proposé de modéliser ces organisations sous forme de graphes bipartites, constitués de sommets de type "région" représentant les éléments observables du chorème et de sommets de type "relation" représentant les liens, décrits par les agronomes, entre ces éléments observables. Nous définissons comme suit les *Graphes d'Organisation Spatiale* :

Définition 4 (gos : Graphe d'Organisation Spatiale) *Un* gos *est un graphe bipartite étiqueté constitué de sommets entité-spatiale (s-entité) et de sommet relation-spatiale (s-relation).*

Définition 5 (s-entité) *Un* s-entité *est un sommet de graphe, instance d'un concept* RÉGION. *Il représente un objet dessiné d'un chorème. Un s-entité peut être :*
- *un objet ponctuel : siège d'exploitation, point d'eau ;*
- *un objet linéaire : route, ruisseau ;*
- *un objet surfacique : champ, pré, ville.*

Définition 6 (s-relation) *Un* s-relation *est un sommet de graphe, instance d'un concept* RELATION. *Un s-relation met en relation des s-entités. Chaque arc d'un s-relation est étiqueté avec le rôle du s-entité dans la relation.*

Voyons comment le modèle proposé peut s'appliquer. Le chorème présenté figure 4.1 décrit l'organisation spatiale d'une exploitation de polyculture-élevage. Deux types de surfaces sont distingués et recouvrent deux activités différentes de l'exploitation. D'une part, les surfaces toujours en herbe, centrées autour du siège de l'exploitation sont destinées à la pâture des animaux - la principale activité de l'exploitation est assurée par la production de lait. D'autre part, des surfaces de culture de céréales destinées pour partie à la vente, l'autre partie assurant le complément alimentaire des animaux. La lecture du chorème permet d'observer différents éléments graphiques représentant des objets de l'exploitation (bâtiments, routes, champs...). La légende qui accompagne le chorème permet d'envisager les propriétés de ces éléments observables. Par contre, la description des organisations spatiales n'est pas évidente à première vue. En effet, les relations ayant du sens entre les objets ne sont pas identifiables immédiatement. Elles sont en fait fournies par la lecture des agronomes qui sont seuls capables d'interpréter le chorème.

Le chorème de la figure 4.1 peut être traduit par le `gos` de la figure 4.2. Ce graphe décrit l'organisation spatiale de l'exploitation modélisée par le chorème. Cette traduction est faite avec l'aide des agronomes qui précisent les relations qui ont du sens pour définir l'organisation spatiale.

Dans le même temps, les agronomes interprètent le chorème et fournissent des explications liées à certaines parties du graphe. C'est ainsi que sont constitués les cas initiaux du système à base de cas. En effet, nous considérons un cas comme étant composé d'un `gos` décrivant une organisation spatiale particulière (le problème) et d'une explication qui lui est associée (la solution). Nous définissons alors des graphes d'organisation spatiale expliqués comme suit :

Définition 7 (gos-e : gos expliqué) *Un* gos-e *est un graphe présentant une organisation spatiale particulière à laquelle est associée une explication qui décrit le fonctionnement, explicité par les agronomes, lié à cette organisation. Un gos-e représente un cas de la base de cas.*

La figure 4.3 présente deux exemples de `gos-e` issus de la description globale de l'exploitation modélisée par le `gos` de la figure 4.2.

4.4 Un modèle des connaissances spatiales

La lecture du chorème permet d'observer différentes entités graphiques, accompagnées d'une légende qui renseigne sur leur identité.

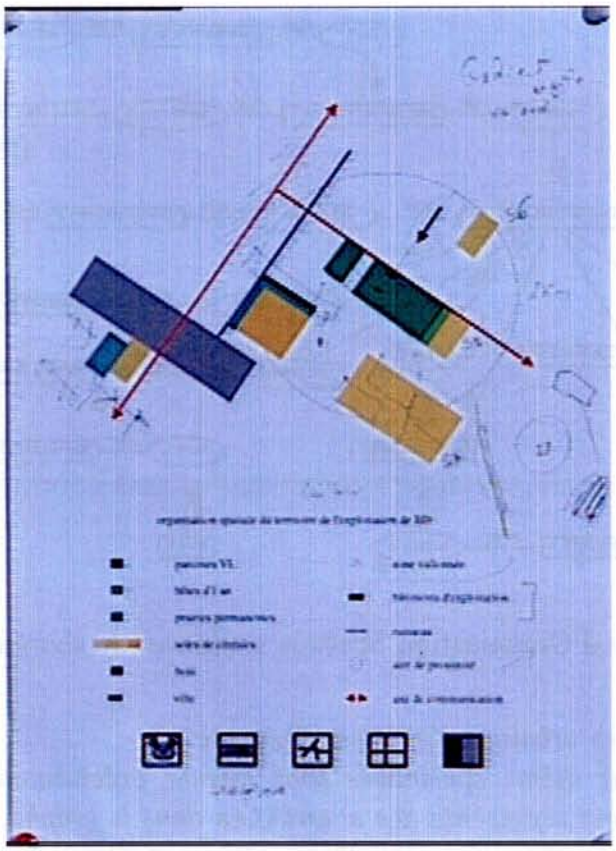


FIG. 4.1 – Chorème représentant l'organisation spatiale d'une exploitation agricole de la région de Vittel. Les traçages au crayon sont le résultat de l'interaction lors de la phase d'acquisition

4.4.1 Les régions

Une fois identifiés les objets présents dans un chorème, on se fie à la légende pour connaître certaines propriétés des entités dessinées. Il s'agit en fait de la définition des concepts du domaine qui seront représentés dans la base de connaissances du système. C'est l'occasion d'approfondir le questionnement des chorèmes. En effet, pour mener à bien ce travail, les informaticiens interrogent les agronomes sur la signification qu'ils donnent aux objets représentés dans les chorèmes, en comparant les différents chorèmes d'exploitations réalisés.

Par étapes successives, nous construisons une base de connaissances dans laquelle les concepts décrivant les s-entités sont spécifiés et organisés. Ces concepts sont construits pas à pas, en définissant leurs attributs à partir de la définition d'un concept de base : RÉGION. Les s-entités sont regroupés sous ce concept qui peut être défini comme suit :

Définition 8 (Concept Région) *Les objets dessinés sur le chorème représentent des régions spatiales de trois types pour lesquels nous définissons les concepts suivants :*

- POINT *: Points d'eau, Bâtiments d'exploitation, etc.*
- LIGNE *: Routes, Ruisseaux, etc.*

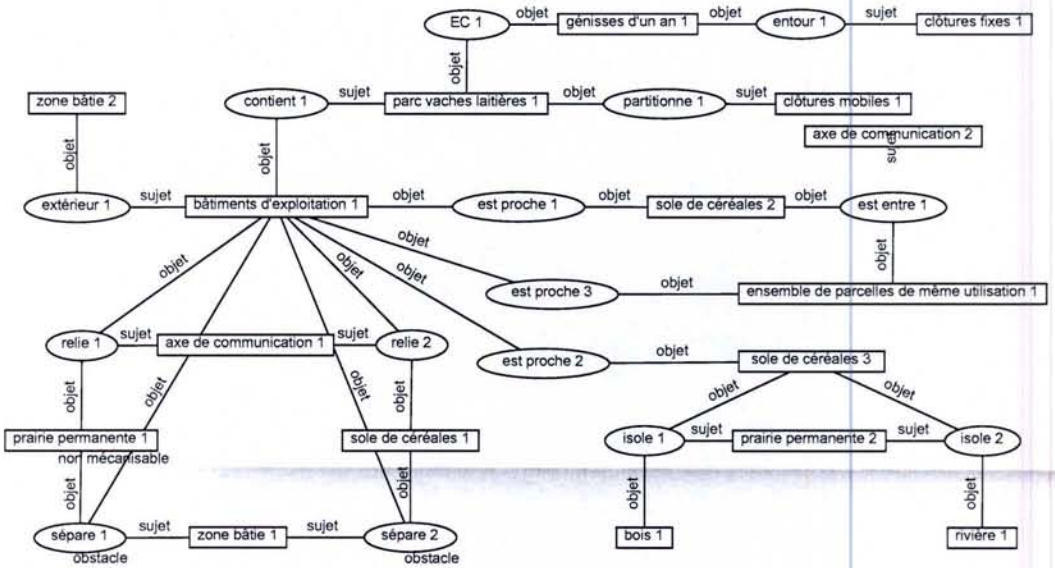


FIG. 4.2 – Graphe d'Organisation Spatiale traduisant le chorème de la figure 4.1.

– SURFACE : *Zones urbaines, Champs, Prés, etc.*

Ces concepts sont eux même spécialisés pour étiquer précisément les différents objets visibles sur le dessin qui représente des **`s-entités`** *dans le graphe.*

Par exemple, sur le chorème présenté figure 4.4 on observe un siège d'exploitation entouré de surfaces en herbe et plus loin, des parcelles de céréales. Pour la surface en herbe, on a évidemment une entité. Pour les céréales, il semble y en avoir plusieurs – plusieurs sommets dans le graphe. En fait après discussion avec les agronomes, il apparaît que l'un et l'autre recouvrent plusieurs entités que l'on considère comme une seule. Nous nous sommes alors mis d'accord sur le fait que les quatre carrés ne représentent pas quatre champs, mais plutôt un certain nombre de parcelles utilisées de façon identique et nous avons défini que ce chorème particulier fait référence à un concept particulier que nous avons noté BLOCS-CÉRÉALES.

Ainsi, chaque concept peut être caractérisé par différents attributs et classé dans une *hiérarchie d'occupations du sol* qui s'est co-construite au fur et à mesure des échanges. Nous avons défini deux hiérarchies de RÉGION pour organiser les concepts manipulés sur les Causses et dans la région de Vittel. Ces deux hiérarchies ont une origine commune mais les termes manipulés pour décrire les objets de ces deux domaines ayant parfois des sens différents, et les deux domaines n'étant pas vraiment comparables, nous avons préféré les séparer. La figure 4.5 présente la hiérarchie des concepts des s-entités pour la région de Vittel.

4.4.2 Les relations

Les chorèmes permettent de représenter des entités spatiales (bâtiments, champs, routes...) qui sont organisées sur le dessin, non pas en fonction de leur positionnement

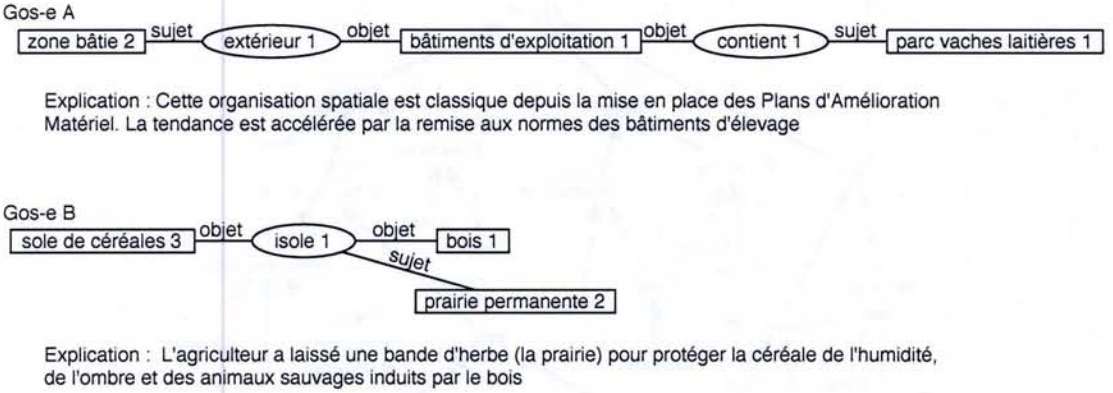


FIG. 4.3 – Exemples d'explications de fonctionnement de l'exploitation agricole en relation avec son organisation spatiale décrite par le graphe de la figure 4.2.

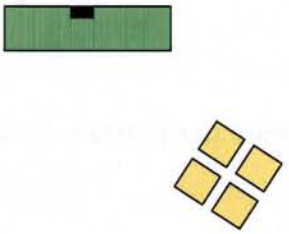


FIG. 4.4 – Exemple de chorème représentant un siège d'exploitation au milieu de surfaces toujours en herbe avec des parcelles de céréales (carrés jaunes) plus éloignées.

topologique, mais en fonction de leurs relations fonctionnelles aux autres entités. Les relations entre les entités ne sont pas directement observables, mais sont déterminées par l'interprétation que fait l'agronome à la lecture du chorème. Après l'identification des entités présentes dans les chorèmes et leur catégorisation, vient le temps d'analyser l'arrangement des objets afin d'identifier leurs relations et de constituer les graphes.

Définition 9 (Concept Relation) *Les chorèmes sont interprétés par les agronomes qui font ressortir certaines relations spatiales et fonctionnelles entre les objets dessinés. Les relations sont organisées dans une hiérarchie dont l'origine est donnée par la description de concept* RELATION.

Ici plus encore, l'expertise des agronomes est indispensable car les relations sont implicites dans le schéma chorématique : seul celui qui l'a dessiné est en mesure de dire si telle relation est pertinente pour rendre compte d'une organisation spatiale.

Par exemple, sur le chorème de la figure 4.1, quelles sont les relations que peuvent entretenir la rivière, la grande parcelle de céréales et la prairie permanente qui est placée entre les deux ? On peut écrire plusieurs graphes dans lesquels les relations n'ont pas la même valeur de pertinence pour les agronomes. Le premier graphe (figure 4.6a) ne rend compte que de l'arrangement des objets du chorème : on constate que le pré est connexe à la rivière et que le pré est connexe au champ. Mais si l'on sait que l'agriculteur a introduit

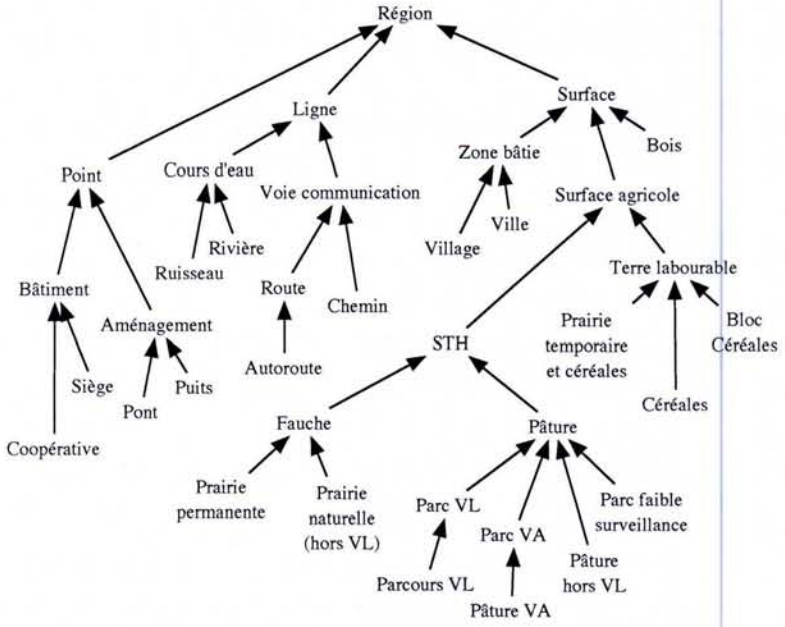


FIG. 4.5 – Hiérarchie des concepts permettant de décrire les éléments du chorème.

un pré entre la rivière et le champ pour faire une zone tampon : on affirme que le pré isole le champ de la rivière. D'où la nécessité de le préciser dans le second graphe, qui est celui retenu par les agronomes (figure 4.6b). On voit ici comment la définition des relations, en interrogeant l'organisation spatiale de l'exploitation représentée dans un chorème, fait apparaître une interprétation du fonctionnement de l'exploitation agricole liée à cette organisation spatiale.

Le choix des relations est parfois délicat, comme nous venons d'en montrer un exemple. En particulier, le choix des relations de proximité ou d'éloignement est important. En effet, une surface peut être considérée comme éloignée si elle se situe à plus d'un kilomètre de l'exploitation si les animaux doivent s'y déplacer, alors qu'elle sera considérée comme proche si c'est une surface sur laquelle n'interviennent que des travaux utilisant le tracteur. Ou encore, une surface peut être considérée comme éloignée, même si elle est physiquement proche, si un obstacle au déplacement existe pour l'atteindre (exemple d'une surface de culture de l'autre côté d'une ville par rapport au siège de l'exploitation, ou encore un pré situé de l'autre côté d'une route à fort trafic).

Suite aux échanges entre agronomes et informaticiens, nous avons organisé les relations manipulées dans une hiérarchie fondée sur les relations topologiques décrites dans la théorie RCC8 [112]. Le treillis des relations topologiques est étendu pour définir des relations fonctionnelles notamment des relations de distance ou des relations ternaires autour de la relation EST-ENTRE. La figure 4.7 présente la hiérarchie de relations telle qu'elle est actuellement définie. Une seule hiérarchie de relations a été définie, car les termes utilisés d'une région à l'autre pour décrire les organisations spatiales recouvrent les mêmes notions.

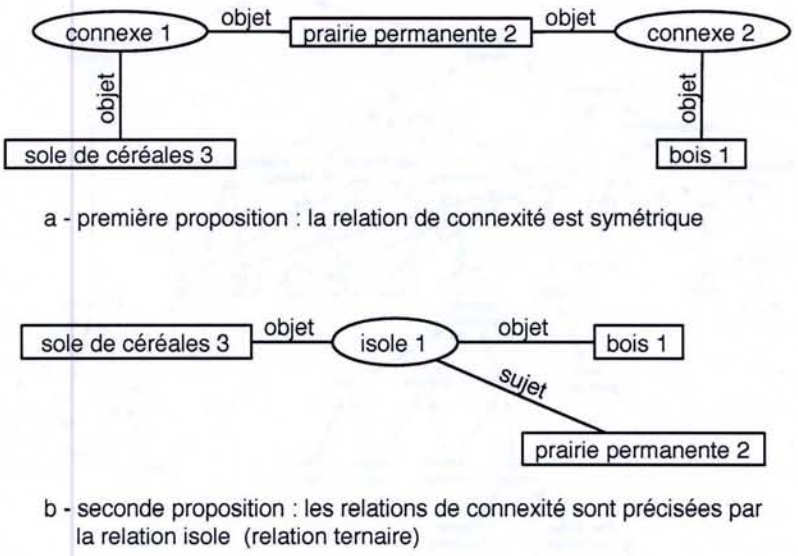


FIG. 4.6 – Deux exemples de traduction de chorème en graphe.

4.4.3 Région ou Relation ?

Parfois, certaines représentations chorématiques posent questions. Il en est ainsi des différents chorèmes présentés figure 4.8, issus de représentations chorématiques décrivant des exploitations du causse Méjan. Le chorème 4.8a représente une surface cultivée, pour la production de céréales ou utilisée comme prairie artificielle (semée), au milieu d'un parcours. La traduction en graphe de ce chorème peut être telle que présentée figure 4.9. Pour le chorème 4.8b, la tentation est grande de donner une traduction en graphe identique alors que de toute évidence les surfaces cultivées ne présentent pas les mêmes caractéristiques. Les situations ne sont sûrement pas comparables. Dans le cas du chorème 4.8c, la première traduction en graphe qui vient à l'esprit est de numéroter les parcelles cultivées et de les relier au parcours par quatre relations identiques. Mais là encore, cette traduction apparaît comme non satisfaisante. En effet, ce n'est pas le fait qu'il y ait quatre champs qui est important ; il peut y en avoir plus. Ce qui veut être mis en avant c'est la présence de plusieurs parties cultivées à l'intérieur d'une surface principalement destinée à la pâture.

En fait, la traduction des chorèmes d'exploitation sous forme de graphes, implique un effort de formalisation tant au niveau des graphes que des chorèmes. L'étape d'acquisition de connaissances tend à faire progresser la connaissance des agronomes. Des connaissances nouvelles ont souvent été le fruit de l'expression de connaissances implicites dont l'explicitation, demandée par les modélisateurs informaticiens, en a révélé le contenu comme nous le présentons au chapitre 7.

Ainsi, nous avons convenus de garder le graphe de la figure 4.9 pour décrire l'organisation spatiale présentée par le chorème 4.8a, alors que nous avons choisi des termes exprimant de nouveaux concepts pour les deux autres chorèmes. Le chorème 4.8b est dénommé sous le terme PRALINE désignant une grande surface de champ entourée de parcours. Le

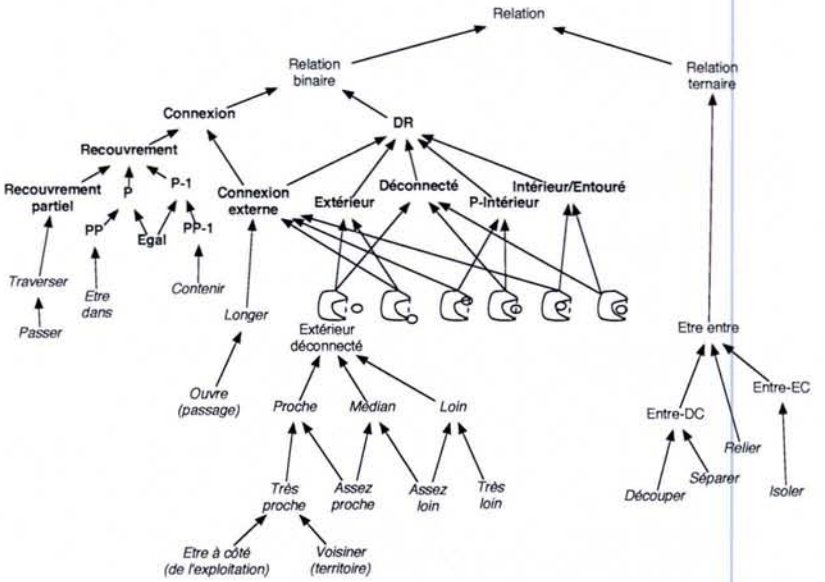


FIG. 4.7 – Hiérarchie des concepts permettant de décrire les relations entre les entités du chorème.

chorème 4.8c est dénommé sous le terme de NOUGAT. Ces deux organisations spatiales sont souvent le fruit d'un choix de conduite des troupeaux par les exploitants, qui donnent lieu à des fonctionnements spécifiques. La PRALINE correspond à une grande proportion de terres cultivables face aux parties en herbe. Ces surfaces sont utilisées pour nourrir des animaux à forts besoins. Le NOUGAT correspond à une plus petite proportion de terres cultivées face aux parties toujours en herbe. Ces surfaces sont en général utilisées pour attirer les brebis sur toute la surface de parcours afin d'en assurer une pâture équilibrée en tous points.

Ainsi, pour le premier cas, le choix de l'utilisation d'une relation est maintenue. Dans les deux autres cas, les concepts spécifiques sont définis et choisis plutôt que des descriptions à l'aide de relations qui rendraient difficiles la distinction des différentes situations avec les graphes.

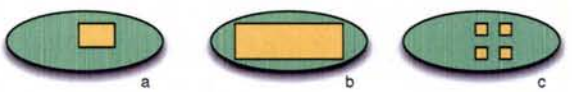


FIG. 4.8 – Exemples de chorèmes rencontrés pour décrire des structures spatiales d'exploitations des causses orientées vers la production de lait de brebis.

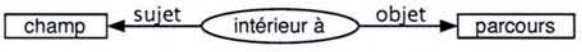


FIG. 4.9 – Traduction sous forme de graphe pour décrire l'organisation spatiale représentée par le chorème de la figure 4.8a.

4.5 Autres connaissances du domaine

Les connaissances du domaine d'ordre spatial permettent d'étiqueter les s-entités et les s-liens à l'aide de concepts pour la modélisation des organisations spatiales agricoles. Ces connaissances sont les premières à avoir été formalisées. Elles sont utilisées lors de la recherche de similarité entre les organisations dans le cadre du RÀPC. Cependant, ces connaissances ne sont pas suffisantes pour envisager le développement d'un système complet. Des connaissances sur les explications et sur les cas doivent être apportées afin de rendre le système opérationnel.

4.5.1 Les explications

La figure 4.3 présente deux exemples d'explications telles qu'elles nous sont fournies par les agronomes pour des situations spatiales particulières.

Le `gos-e A` modélise une organisation spatiale où le parc à vaches laitières entoure les bâtiments d'exploitation, qui sont hors du village. Cette organisation a pour but de minimiser le déplacement des animaux tout en offrant un espace suffisant pour le troupeau. Cette structure est définie par les agronomes comme étant très classique surtout depuis l'application de nouvelles réglementations, et des aides qui les ont accompagnées pour leur mise en place. Ainsi, les bâtiments ont été sortis de la commune pour permettre leur agrandissement et une meilleure hygiène.

Ce que nous demandons aux agronomes, c'est d'expliciter les explications qu'ils associent aux organisations spatiales. La question posée est : « Etant donné la structure spatiale représentée par le graphe, quel est le rôle de chacun des sommets dans l'explication ? » Nous cherchons à obtenir une bijection entre la structure du graphe et la structure de l'explication.

L'explication ainsi retravaillée devient alors : « Les bâtiments d'exploitation sont au milieu des parcs pour limiter le déplacement des animaux et ils sont placés à l'extérieur du village pour permettre des constructions de plus grande taille, que celles possibles dans le village. » Les différents termes de l'explication peuvent être rattachés aux différents sommets du graphe.

Le `gos-e B` décrit une organisation spatiale particulière : l'étiquette `isole 1` rend compte d'une relation spatiale qui possède aussi un aspect fonctionnel. L'entité `prairie permanente 2` est isolante, tandis que l'entité `sole de céréales 3` et l'entité `bois-1` sont isolées. L'explication associée à ce graphe est : « L'agriculteur a placé une prairie pour isoler les céréales de l'humidité, des animaux et de l'ombre provenant du bois ».

Les explications sont actuellement des phrases associées à des `gos` pour définir des `cas` ou `gos-e`. Mais, dans le but de permettre l'adaptation, il reste un travail d'acquisition de connaissances pour les modéliser plus avant. Pour cela, nous envisageons de les décrire en reliant les termes des phrases aux différents sommets (RÉGION ou RELATION) du `gos` associé auxquels ils font référence. De cette façon, nous pourrons transformer les termes d'une explication d'un cas source en fonction du chemin de similarité qui a permis de rendre comparable les graphes source et cible. Cette structuration des explications permettra de les classer les unes par rapport aux autres dans une hiérarchie d'explication initiée par le concept EXPLICATION.

Définition 10 (Concept Explication) *Les explications sont données par les agronomes. Elles lient fonctionnement et organisation spatiale des exploitations. Les explications restent à être formalisées. Elles sont actuellement enregistrées sous forme de texte brut.*

Ainsi, l'explication du `gos-e B` peut être adaptée à l'organisation spatiale très similaire, visible sur le chorème d'exploitation de la figure 4.2, où la bande d'herbe est située entre la parcelle de céréales et la rivière. Dans le `gos`, le s-entité `bois-1` est alors remplacé par `rivière-1`. Dans l'explication, les termes faisant référence au bois sont remplacés par des termes faisant référence à la rivière. On obtient ainsi le `gos-e C` de la figure 4.10.

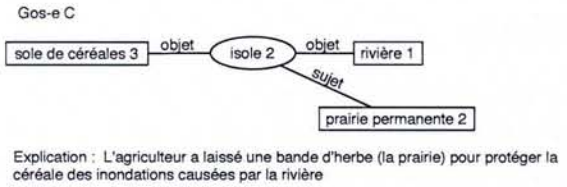


FIG. 4.10 – Exemple d'adaptation d'une explication.

En revanche, si la situation concerne une parcelle de céréales qui longe un bois ou une rivière que pourra-t-on en dire ? L'explication pourrait être : « L'agriculteur ne considère pas comme un problème la présence d'une rivière en bordure de sa parcelle de céréales ». Comme on le constate ici, l'adaptation de l'explication nécessite des outils de traitement automatique du langage que nous n'avons pas abordés dans le cadre de nos travaux. La manipulation de la négation dans les explications pose de nombreux problèmes qui seront à résoudre pour la mise en place d'une adaptation de qualité.

4.5.2 Les points de vue

Dans l'exemple du cas, tiré d'une exploitation du causse Méjan et présenté figure 4.11, le parcours est rendu intéressant pour l'exploitant par la présence du point d'eau qui attire les brebis et permet ainsi son pâturage.

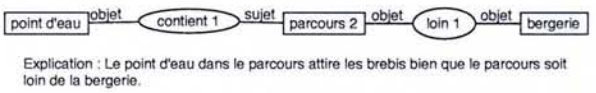


FIG. 4.11 – Exemple de cas pour lequel un point de vue d'attrait des brebis est présent.

Il arrive que l'explication soit similaire dans le cas où ce n'est plus un point d'eau qui est contenu dans le parcours, mais une petite parcelle cultivée. Dans ce cas, la parcelle et le point d'eau jouent le même rôle, i.e. attirer les brebis dans le parcours pour le pâturer. Or selon la hiérarchie des concepts, un champ (une surface cultivée) n'est pas vraiment comparable à un point d'eau (un aménagement ponctuel). Ceci tient au fait que la hiérarchie est construite selon un point de vue assez naturel, une première entrée pour décrire les objets d'une exploitation, à savoir l'occupation du sol. Ici ce n'est plus

ce point de vue qui prévaut, mais bien le point de vue de l'attrait des animaux vers une portion du territoire. On souhaite donc pouvoir rendre comparable dans cette situation le point d'eau, le champ, mais aussi un abri ou un bois qui offrent des éléments de sécurité pour les animaux. En plus du point de vue de déplacement du troupeau, nous avons observé les points de vue : de surveillance, d'alimentation, de contraintes de déplacement occasionné par des routes ou encore des agglomérations. Pour ces points de vue, on désire donc établir des passerelles entre différents concepts de la hiérarchie d'occupation du sol pour les regrouper. Dans ces points de vue, les concepts regroupés sont considérés comme similaires.

4.5.3 Les échelles

Les graphes sont décrits à plusieurs niveaux de granularité, à différentes ECHELLES et donc les cas aussi. Trois niveaux ont été observés durant la phase d'acquisition de connaissances. Ce sont les niveaux d'organisation à l'échelle du *territoire*, de l'*exploitation* ou de la *parcelle*.

L'échelle du territoire, correspond à l'observation des relations de l'exploitation avec son environnement : les routes que relient le siège, la position par rapport au village... Ensuite, vient le niveau de l'exploitation dans lequel les cas traitent principalement de la circulation des animaux ou des travaux de production (céréales, viande, lait). Enfin, au niveau de la parcelle, les cas décrivent les aménagements spécifiques réalisés par l'exploitant ou encore les problèmes de voisinage d'une parcelle : un bois à côté d'un pré présente le risque de voir des ligneux s'installer sur la parcelle, c'est l'embroussaillement, problème qui est étudié dans le cadre des recherches sur les causses. Aussi on ne comparera pas des cas à des ECHELLES différentes.

4.6 Conclusion

La transformation des chorèmes en graphes répond à deux objectifs. Le premier est lié à la construction du système de RÀPC : il s'agit de constituer progressivement une base de cas et une base de connaissances sur le domaine considéré. Le second, au moyen de la reformulation des connaissances représentées dans les chorèmes, est de faire progresser les connaissances des agronomes. L'utilisation des graphes pour modéliser les chorèmes démontre que :

- Les objets du chorème peuvent être modélisés par les entités d'un graphe,
- Les caractéristiques de ces objets (formes et légendes associées) peuvent être décrites par des attributs associés aux entités du graphe,
- Leur arrangement peut être modélisé par des relations d'un graphe.

L'interaction entre agronomes et informaticiens a poussé les premiers à reformuler les informations contenues dans les chorèmes, à expliciter le raisonnement à la base de leur dessin. Enfin, un travail sur l'analyse des raisonnements des agronomes dans la comparaison des situations entre différentes exploitations, a permis de déterminer des règles d'adaptation des explications en fonction des différences au niveau des organisations spatiales.

L'ensemble de ces phases permet de déterminer les éléments nécessaires au développement d'un système à base de connaissances.

Cette démarche illustre l'intérêt de la combinaison d'outils et de la complémentarité de disciplines pour modéliser des phénomènes complexes. Elle ne s'appuie pas uniquement sur l'interaction entre des chercheurs de disciplines différentes ou sur la mobilisation conjointe d'outils. Elle repose sur une association continue entre des points de vue sur un territoire (les règles de fonctionnement des exploitations en sont l'expression) et des propriétés de l'espace (révélées par la structure de l'exploitation) pour aboutir à un modèle qui intègre le spatial et le fonctionnel, et qui de ce fait renvoie aux stratégies des exploitants agricoles, primordiales dans la transformation des territoires ruraux.

Au terme de cette phase d'acquisition de connaissances, nous avons modélisé la description de cinq exploitations en polyculture-élevage laitier sur la région Lorraine et quatre exploitations des causses en élevage ovin-viande. Nous avons produit deux hiérarchies de concepts d'occupation du sol pour chacune des régions étudiées. Nous avons aussi défini une hiérarchie des relations utilisées pour décrire les organisations spatiales. Enfin, nous avons défini une série de concepts pour représenter les diverses autres connaissances que devra manipuler le système ROSA. L'ensemble de ces hiérarchies n'est pas figé et évolue et peut être remis en question en fonction des connaissances nouvelles apportées par l'étude de nouvelles exploitations ou/et de nouvelles régions.

Le chapitre suivant décrit comment sont représentées l'ensemble de ces connaissances à l'aide du langage de représentation que nous avons choisi i.e. les logiques de descriptions. Puis, nous nous intéressons à leur utilisation dans les mécanismes de raisonnement mis en place dans notre système.

S.C.D. - U.H.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

Chapitre 5

Représentation des connaissances et des raisonnements

Sommaire

5.1 Sélection du langage de représentation 84
5.2 Connaissances terminologiques 84
5.2.1 Concepts 84
5.2.2 Rôles 85
5.3 Connaissances factuelles 86
5.3.1 Sommets 87
5.3.2 Graphes 88
5.3.3 Cas 90
5.4 Le RàPC dans ROSA 90
5.4.1 Principe 92
5.4.2 Classification dure 93
5.4.3 Une classification « semi-dure » 93
5.4.4 Transformation des graphes 95
5.4.5 Classification élastique 96
5.4.6 Pistes pour l'adaptation 97

Nous présentons dans ce chapitre comment nous avons utilisé les logiques de descriptions pour représenter les connaissances du domaine mais aussi les connaissances sur les graphes et les cas. Ensuite, nous présentons comment nous avons utilisé les mécanismes de raisonnement fournis par les logiques de descriptions dans le raisonnement à partir de cas et plus particulièrement pour les deux principales opérations que sont la remémoration et l'adaptation.

5.1 Sélection du langage de représentation

Pour représenter les connaissances acquises et définies au chapitre précédent, plusieurs formalismes sont envisageables comme nous l'avons vu dans le chapitre 1. Lors des travaux précurseurs de cette thèse, le système YAFOOL fut utilisé. Il n'est malheureusement plus développé et soutenu. Des comparaisons avec les LD avaient été entreprises alors et avaient permis d'étendre YAFOOL avec de nouveaux opérateurs, donnant à ce langage l'expressivité des systèmes de LD existants [64].

Des travaux sur le RÀPC à l'aide de LD ont été entrepris [115]. Les LD sont particulièrement efficaces pour des raisonnements de classification en milieu hiérarchique qui correspondent à nos besoins. Notre manipulation des graphes aurait pu nous faire choisir le formalisme des graphes conceptuels [97]. Cependant ce formalisme n'a pas, à notre connaissance, été utilisé dans le cadre du RÀPC, et les quelques expériences réalisées ont montré l'intérêt mais aussi les limites d'outils génériques tels que Corese ou Cogitant. C'est donc assez naturellement que nous nous sommes intéressés aux LD et notamment à RACER qui offre un très bon support pour la manipulation de concepts et le calcul de subsomption et aussi un support des individus à l'aide de la `ABox`.

5.2 Connaissances terminologiques

Les connaissances du domaine sont enregistrées dans la base terminologique (`TBox`) et organisées selon une hiérarchie des concepts $\mathcal{H}_C$ et une hiérarchie des rôles $\mathcal{H}_R$.

5.2.1 Concepts

Pour les besoins de notre application, les premiers concepts primitifs de $\mathcal{H}_C$ (figure 5.1) sont les concepts suivants :

- SOMMET. Il représente les sommets des graphes d'organisation spatiale. Un sommet peut représenter une entité spatiale de type RÉGION ou une relation entre deux entités de type RELATION. Ces deux concepts RÉGION et RELATION initient les deux hiérarchies $\mathcal{H}_{CE}$ (figure 4.5) et $\mathcal{H}_{CR}$ (figure 4.7).
- GOS. Une instance de type GOS est définie par une liste de sommets dont le type est SOMMET. La structure est assurée via les rôles inclus dans la définition des RELATION.
- EXPLICATION. Ce concept représente les explications associées aux organisations spatiales particulières pour composer les cas.

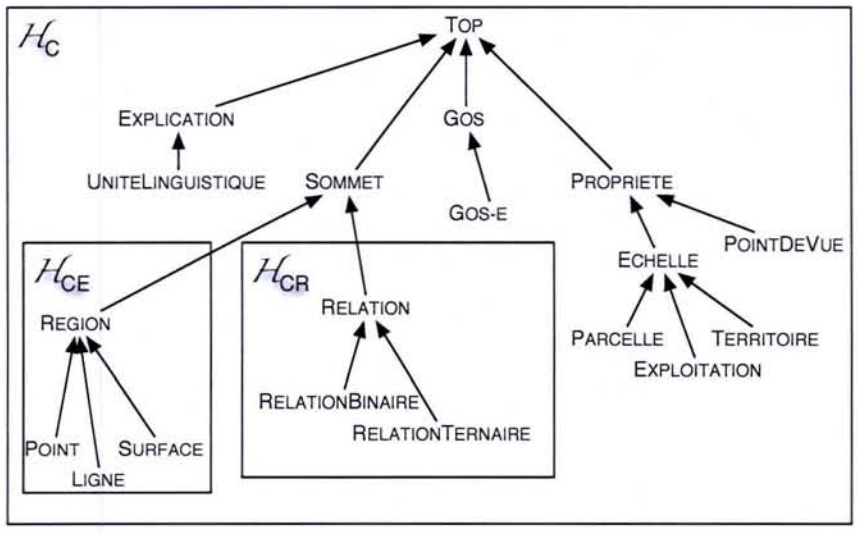


FIG. 5.1 – Hiérarchie des concepts manipulés dans ROSA.

- CAS. Ce concept représente les cas de la base de cas. Il est dérivé du concept GOS auquel une EXPLICATION est associée.
- PROPRIÉTÉ. Permet de préciser les propriétés de différents objets et notamment des propriétés associées aux sommets `s-entité` et `s-relation`.
- ECHELLE. Ce concept est une des premières propriétés dont nous avons eu besoin pour préciser les cas.
- POINTDEVUE. Les points de vue définissent une propriété des cas. Ils s'apparentent aux passerelles proposées dans TROEPS. Un point de vue rend comparable des concepts apparemment non comparables selon la hiérarchie des concepts du domaine organisée par le point de vue implicite d'occupation du sol.

Ces premiers concepts primitifs sont introduits dans le langage de RACER par le constructeur *implies*. La description du concept primitif SOMMET s'écrit comme suit :

```
(implies SOMMET TOP).
```

Ils sont ensuite spécialisés, en fonction des besoins de représentation de connaissances du domaine, par d'autres concepts primitifs ou des concepts définis. Par exemple, le concept RELATION est défini comme un concept SOMMET qui est relié à des RÉGION à l'aide du rôle *arc*. Il s'écrit en RACER :

```
(define RELATION (and SOMMET (some arc RÉGION))).
```

Le tableau 5.1 donne la définition dans le formalisme des LD des premiers concepts manipulés dans notre système ROSA.

5.2.2 Rôles

La hiérarchie des rôles $\mathcal{H}_R$ est présentée figure 5.2. Le rôle *possèdeSommet* permet de relier les sommets à leur graphe. Le domaine de ce rôle est donc le concept GOS alors

SOMMET $\sqsubseteq$ TOP
GOS $\sqsubseteq$ TOP
EXPLICATION $\sqsubseteq$ TOP
PROPRIÉTÉ $\sqsubseteq$ TOP
POINTDEVUE $\sqsubseteq$ PROPRIÉTÉ
ECHELLE $\sqsubseteq$ PROPRIÉTÉ
CAS $\sqsubseteq$ GOS $\sqcap$ ($\exists$*explique* $\cdot$ EXPLICATION)
RÉGION $\sqsubseteq$ SOMMET $\sqcap \neg$RELATION
RELATION $\doteq$ SOMMET $\sqcap$ ($\exists$*arc* $\cdot$ RÉGION)
RELATIONBINAIRE $\doteq$ RELATION $\sqcap$ ($= 2$*arc* $\cdot$ RÉGION)
RELATIONTERNAIRE $\doteq$ RELATION $\sqcap$ ($= 3$*arc* $\cdot$ RÉGION)

TAB. 5.1 – Premiers concepts de la hiérarchie $\mathcal{H}_C$. RÉGION initie la hiérarchie $\mathcal{H}_{C\mathcal{E}}$ et RELATION la hiérarchie $\mathcal{H}_{C\mathcal{R}}$.

que le co-domaine prend ses valeurs dans SOMMET. Le rôle *arc* permet de définir les arcs des graphes. Ces arcs sont orientés des sommets de type RELATION (domaine) vers les sommets de type RÉGION (co-domaine) ce qui garantit la bi-partition des graphes. Un *arc* peut être de deux types, *sujet* ou *objet*, en fonction du rôle de l'entité spatiale dans la relation.

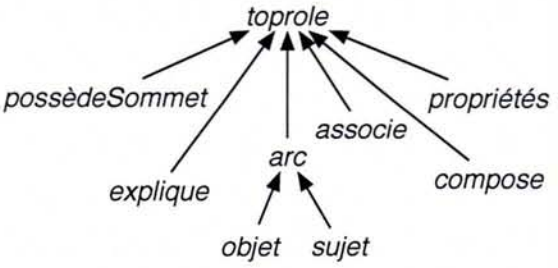


FIG. 5.2 – Hiérarchie des rôles $\mathcal{H}_R$ utilisés pour la représentation des connaissances dans notre système.

Le rôle *arc* est introduit dans le langage RACER par l'expression :

```
(arc :domain RELATION :range RÉGION),
```

alors que le rôle *objet* est introduit par l'expression :

```
(objet :parent arc).
```

5.3 Connaissances factuelles

Les connaissances factuelles, enregistrées dans une ABox servent à définir les individus (i.e. les connaissances sur les exploitations), leurs propriétés et leurs relations. Des individus peuvent représenter des entités spatiales du chorème mais aussi des relations entre

$$
\begin{array}{c}
\textit{arc} \sqsubseteq \textit{toprole} \\
\textit{propriétés} \sqsubseteq \textit{toprole} \\
\textit{objet} \sqsubseteq \textit{arc} \\
\textit{sujet} \sqsubseteq \textit{arc} \\
\textit{explique} \sqsubseteq \textit{toprole} \\
\textit{possedeSommet} \sqsubseteq \textit{toprole} \sqcap \textit{sommetDe}^{-1} \\
\textit{compose} \sqsubseteq \textit{toprole} \\
\textit{associe} \sqsubseteq \textit{toprole}
\end{array}
$$

TAB. 5.2 – Définition des rôles dans la hiérarchie $\mathcal{H}_R$.

ces entités ; ce sont les sommets des graphes qu'ils constituent. Un graphe représente les connaissances d'ordre spatial d'une exploitation alors que les cas définissent des organisations spatio-fonctionnelles. Nous nous servirons du graphe de la figure 5.3 pour illustrer comment nous représentons les connaissances sur une exploitation dans le cadre des LD et du système RACER en particulier.

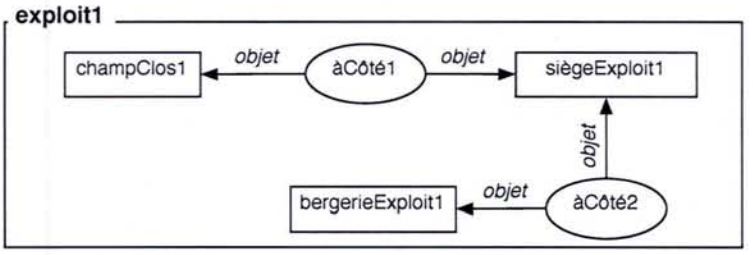


FIG. 5.3 – Exemple de graphe décrivant des connaissances sur une exploitation.

5.3.1 Sommets

Chaque sommet est défini par un nom et instancié avec un type de la hiérarchie $\mathcal{H}_{\mathcal{CE}}$ pour les `s-entité` ou de la hiérarchie $\mathcal{H}_{\mathcal{CR}}$ pour les `s-relations`. Par exemple un premier champ clôturé est déclaré comme étant un champ, ce qui s'écrit en RACER :

```
(instance champClos1 CHAMP).
```

Ce `champClos1` est situé à côté d'une autre entité spatiale. Nous réifions alors la relation ETRE-A-COTÉ comme nous l'avons fait pour le champ clos :

```
(instance aCôté1 ETRE-A-CÔTÉ).
```

Après avoir défini des instances, il faut déclarer les relations entre elles : les arcs, qui définissent la structure d'un graphe, sont ajoutés. Un arc est représenté par un `rôle` des LD en fonction du rôle du `s-entité` dans la relation. Si la relation est symétrique, le rôle *objet* sera choisi pour chaque arc de la relation, sinon nous aurons un rôle *sujet* et un rôle *objet* pour une relation binaire ou un *sujet* et deux *objet* pour une relation ternaire. Pour notre exemple nous écrivons en RACER :

```
(related aCôté1 champClos1 objet).
```

Ce qui signifie que le `champClos1` est relié à la relation `aCôté1` en jouant le rôle d'*objet* dans la relation.

Pour les `s-entités` comme pour les `s-relations`, en plus de connaître leurs relations, nous souhaitons préciser leur propriétés. En LD, un individu ne peut être mis en relation qu'avec un autre individu. Dans ces conditions, comment définir qu'un champ, par exemple, est clôturé ? Une première approche consiste à définir un concept CLÔTURE et à l'instancier avec une instance générique (prototypique). Il est donc nécessaire de construire autant d'instances génériques qu'il peut y avoir de propriétés à définir sur les différents `s-entités` et `s-relations` des graphes. Cette solution n'est pas très élégante et rend la représentation des propriétés associées aux individus assez complexe. Une autre solution proposée par C. A. Welty [124] consiste à définir une instance qui hérite de plusieurs concepts de PROPRIÉTÉ et de lier cette instance à l'objet qui possède toutes ces propriétés. Ce qui s'écrit dans le formalisme de RACER :

```
(instance prop-champClos1 CAILLOUTEUX)
(instance prop-champClos1 CLOS)
(related champClos1 prop-champClos1 propriétés).
```

Où CAILLOUTEUX et CLOS sont des sous concepts de PROPRIÉTÉ.

5.3.2 Graphes

Le graphe d'organisation spatiale, constitué de sommets `s-entité` et `s-relation`, représente la principale structure de connaissance manipulée dans notre système. Chaque exploitation est décrite par un `gos` associé à son chorème dont il est une traduction en terme informatique. Le `gos` d'une exploitation comme les `s-entités` et les `s-relations` sont des individus au sens des LD. Chaque graphe est donc défini par son nom et instancié avec le type GOS :

```
(instance exploit1 Gos).
```

La représentation d'un graphe est ainsi simplement réalisée dans le cadre des LD, mais le lien entre l'instance d'un graphe et sa composition n'est pas encore assurée. Pour ce faire, on peut déclarer le graphe comme étant constitué d'une liste de sommets par l'intermédiaire du rôle *possèdeSommet* qui déclare l'appartenance de chaque sommet au graphe. Ainsi notre champ clôturé est ajouté au graphe comme suit :

```
(related exploit1 champClos1 possèdeSommet).
```

Nous avons donc réalisé la représentation des graphes dans le cadre des LD à l'aide d'une énumération plate des sommets associés à un graphe et en encodant indirectement la liste des arcs grâce à la réification des relations. Seule la liste des sommets est enregistrée avec les `gos` ; les arcs sont retrouvés si nécessaire au travers des sommets de type RELATION auxquels ils sont attachés.

La représentation complète de l'exemple de la figure 5.3 est décrit dans le tableau 5.3. Il est à noter que les graphes décrivant les exploitations sont beaucoup plus gros (cf. figure 4.2). Ici nous avons plutôt la représentation d'un « graphe de cas ».

```
(instance siègeExploit1 SIÈGE)
(instance bergerieExploit1 BERGERIE)

(instance champClot1 CHAMP)
(instance prop-champClos1 CAILLOUTEUX)
(instance prop-champClos1 CLOS)
(related champClos1 prop-champClos1 propriétés)

(instance aCôté1 ETRE-A-CÔTÉ)
(related aCôté1 siègeExploit1 objet)
(related aCôté1 bergerieExploit1 objet)

(instance aCôté2 ETRE-A-CÔTÉ)
(related aCôté2 siègeExploit1 objet)
(related aCôté2 champClos1 objet)

(instance exploit1 GOS)
(related exploit1 siègeExploit1 possèdeSommet)
(related exploit1 bergerieExploit1 possèdeSommet)
(related exploit1 champClos1 possèdeSommet)
(related exploit1 aCôté1 possèdeSommet)
(related exploit1 aCôté2 possèdeSommet)
```

TAB. 5.3 – Liste des commandes permettant la représentation du graphe de la figure 5.3 avec RACER.

5.3.3 Cas

Un cas est constitué d'un sous-graphe d'un graphe décrivant une exploitation auquel est associée une explication. Les sommets sont sélectionnés par leur nom. La structure d'une explication est définie par une liste d'unités linguistiques qui constituent la phrase explicative du cas. Chaque unité linguistique faisant référence à un des sommets du `gos` est alors associée au sommet correspondant. Les autres unités linguistiques sont simplement énoncées sans référence à un sommet. L'exemple ci-dessous présente le cas associé à l'organisation spatiale décrite dans le tableau 5.3. Le cas est défini à l'échelle du territoire de l'exploitation. L'explication donnée est :

> *« L'agriculteur utilise ce champ à fort potentiel (ressource),clôturé (sans garde) proche du siège (surveillance) et proche de la bergerie(déplacements limités) pour y mettre les animaux à très forts besoins. »*

Cet énoncé se compose de huit unités linguistiques de la façon suivante :
- `ul1` : "L'agriculteur utilise"
- `ul2` : "ce champ à fort potentiel (ressource), clôturé (sans garde)"
- `ul3` : "proche"
- `ul4` : "du siège (surveillance)"
- `ul5` : "et"
- `ul6` : "proche"
- `ul7` : "de la bergerie (déplacements limités)"
- `ul8` : "pour y mettre les animaux à très forts besoins."

Les unités linguistiques 2,3,4,6 et 8 font respectivement référence aux sommets `champClos1`, `aCôté2`, `siègeExploit1`, `aCôté1`, `bergerieExploit1`. La description sous RACER est donnée dans le tableau 5.4. La catégorisation des unités linguistiques, ainsi que leur enchaînement pour construire l'explication reste à développer. En effet, la manipulation d'ensembles ordonnés en LD n'est pas actuellement possible. En ce qui concerne la catégorisation et la manipulation des explications en vue de les adapter pour le RÀPC, il faudra au minimum dans la base de connaissances définir les usages du territoire, mais aussi des outils de traitement automatique du langage que nous n'avons pas abordés dans ce travail. Le concept unique d'UNITÉLINGUISTIQUE devra être spécialisé, comme nous avons spécialisé les hiérarchies $\mathcal{H}_{\mathcal{CE}}$ et $\mathcal{H}_{\mathcal{CR}}$.

5.4 Le RàPC dans ROSA

Le cycle de raisonnement dans ROSA s'organise autour des deux principales opérations que sont la remémoration et l'adaptation. La remémoration est fondé sur les principes de classification de classification dure et élastique [72]. La base de cas est organisée par une structure hiérarchique des index de cas telle que présentée figure 5.4. Cette hiérarchie est fondée sur des graphes concepts sous lesquels sont classés les cas. La relation de généralité entre graphes est donnée par la définition suivante :

Définition 11 *Un graphe $G1$ est plus spécifique qu'un graphe $G2$ s'il contient un sous-graphe $G'1$ isomorphe au graphe $G2$ tel que les concepts associés aux sommets de $G'1$ sont plus spécifiques ou bien identiques à ceux du graphe $G2$.*

```
(instance cas1 CAS)
(instance prop-cas1 TERRITOIRE
(instance explic1 EXPLICATION)
(instance ul1 UNITÉLINGUISTIQUE)
(instance ul2 UNITÉLINGUISTIQUE)
(instance ul3 UNITÉLINGUISTIQUE)
(instance ul4 UNITÉLINGUISTIQUE)
(instance ul5 UNITÉLINGUISTIQUE)
(instance ul6 UNITÉLINGUISTIQUE)
(instance ul7 UNITÉLINGUISTIQUE)
(instance ul8 UNITÉLINGUISTIQUE)

(related cas1 prop-cas1 propriétés
(related cas1 explic1 propriétés
(related explic1 ul1 compose)
(related explic1 ul2 compose)
(related explic1 ul3 compose)
(related explic1 ul4 compose)
(related explic1 ul5 compose)
(related explic1 ul6 compose)
(related explic1 ul7 compose)
(related explic1 ul8 compose)

(related cas1 siègeExploit1 possèdeSommet)
(related cas1 bergerieExploit1 possèdeSommet)
(related cas1 champCôt1 possèdeSommet)
(related cas1 aCôté1 possèdeSommet)
(related cas1 aCôté2 possèdeSommet)

(related ul2 champClos1 associe)
(related ul3 aCôté2 associe)
(related ul4 siègeExploit1 associe)
(related ul6 aCôté1 associe)
(related ul8 bergerieExploit1 associe)
```

TAB. 5.4 – Représentation en RACER d'un cas constitué de sommets d'un gos et d'une explication associée aux sommets du gos.

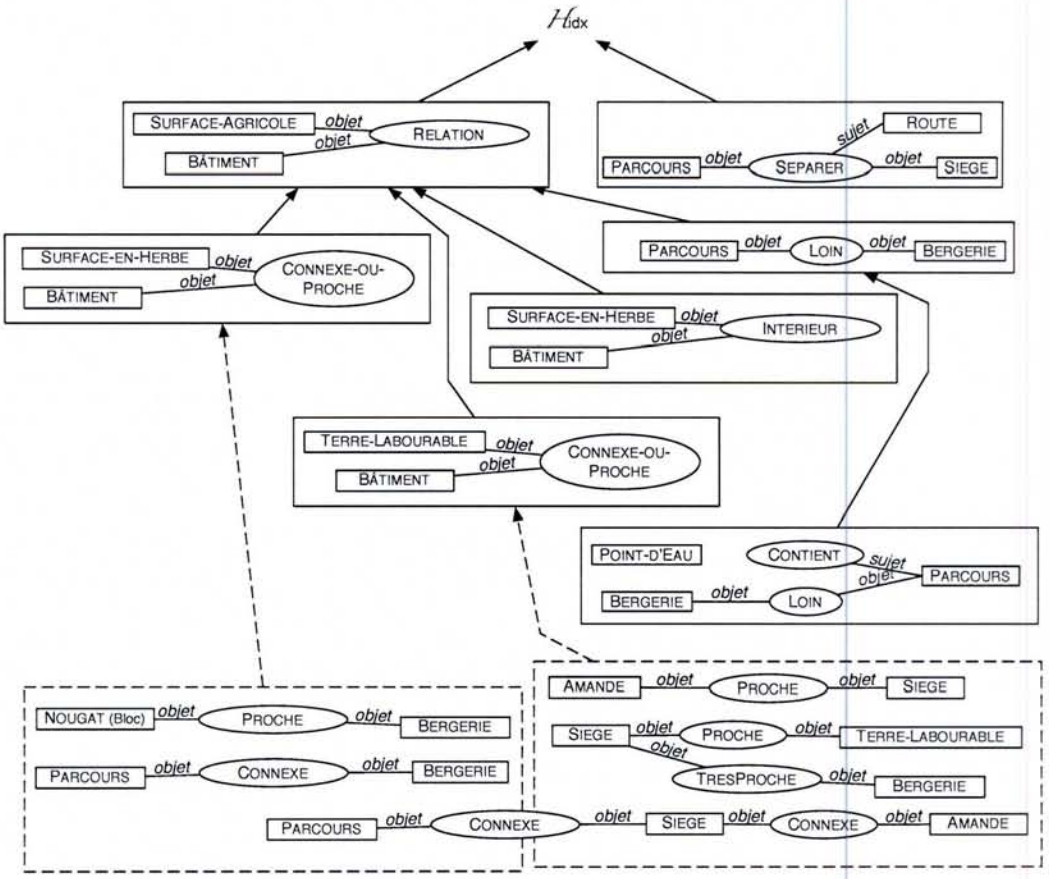


FIG. 5.4 – Hiérarchie partielle des index de cas (les explications ne sont pas représentées).

5.4.1 Principe

L'opération de remémoration est fondée sur la classification d'un graphe d'organisation spatiale d'une exploitation cible `org(cible)` dans la hiérarchie des index de cas $\mathcal{H}_{idx}$. Cette classification permet de déterminer un ou plusieurs cas sources qui peuvent être réutilisés dans le but d'expliquer le fonctionnement de l'exploitation après une phase d'adaptation des explications au nouveau contexte. Un ou plusieurs cas sources peuvent être retrouvés qui sont associés à une ou plusieurs parties du graphe cible à expliquer. Lorsque plusieurs cas sources sont associés à une même partie du graphe cible, alors une sélection du meilleur cas source est faite en fonction de critères de préférence donnés. Ce type de classification est désigné sous le nom de *classification dure*. Si certaines parties du graphe cible ne sont associées à aucun cas source alors un processus de *classification élastique* peut être activé. Ce processus consiste à modifier le graphe `org(cible)`, selon certaines *règles de transformation* afin de produire plus d'explications sur l'organisation spatiale cible étudiée.

5.4.2 Classification dure

Dans le cas du processus de classification dure, le chemin de similarité est défini de la façon suivante :

$$\texttt{org(source)} \sqsubseteq \texttt{idx(org(source))} \sqsupseteq \texttt{org(cible)}.$$

La relation $\texttt{org(source)} \sqsubseteq \texttt{idx(org(source))}$ signifie que l'index associé au cas source, i.e. `idx(org(source))`, est plus général que la description de l'organisation `org(source)` ce qui doit être toujours le cas. La relation $\texttt{idx(org(source))} \sqsupseteq \texttt{org(cible)}$ signifie que l'index du cas source est plus général que le cas cible. Nous considérons que le cas cible est son propre index comme proposé dans [74, 75]. Le chemin de similarité ne prend de sens que si `idx(org(source))` est différent de $\top$ dans la hiérarchie $\mathcal{H}_{idx}$.

Dans l'exemple présenté figure 5.5, le graphe d'organisation spatiale `org(cible)` est classé dans la hiérarchie des index $\mathcal{H}_{idx}$. Le graphe `index` est trouvé comme étant un des subsumants les plus spécifiques de `org(cible)` dans la hiérarchie qui est appariée aux sommets (`parc2`, `à côté`, `bergerie`) de `org(cible)`. `org(cible)` est alors comparé aux instances de `index1` de la base de cas. Le chemin de similarité est calculé en comparant des sommets de `org(cible)` à ceux des instances $\texttt{org(source}_x\texttt{)}$. Dans l'exemple de la figure 5.5, deux chemins de similarité sont évalués comme suit :

- De `org(cible)` à `org(source11)`, une première opération d'instanciation des sommets de chaque graphe permet d'obtenir leur graphe concept. Puis ces graphes sont comparés à `index1` grâce aux opérations de généralisation et de spécialisation. On obtient ainsi : PARC $\sqsubseteq$ STH et STH $\sqsupseteq$ PARCOURS, soit deux transformations (une généralisation et une spécialisation).
- De `org(cible)` à `org(source12)`, la première opération d'instanciation des sommets de chaque graphe (rappelons que `à-côté` est une instance de TRÈS-PROCHE) permet de les comparer à `index1`. Deux généralisations sont appliquées : TRÈS-PROCHE vers PROCHE et PARC en STH, puis une spécialisation de STH en PARCOURS. On obtient alors PARC $\sqsubseteq$ STH et STH $\sqsupseteq$ PARCOURS et TRÈS-PROCHE $\sqsubseteq$ PROCHE soit trois transformations.

5.4.3 Une classification « semi-dure »

Au début de nos travaux, le faible nombre de cas à notre disposition ne nous permettait pas d'avoir une hiérarchie d'index intéressante. Aussi, nous avons envisagé de construire à la volée des graphes de concepts index représentant les plus petits subsumants communs entre les graphes de la base de cas et le cas cible donné en entrée du système. Cette solution permet d'apprendre de nouveau cas au "voisinage" plus ou moins proche des cas existants et offre la possibilité de définir des index susceptibles d'être suffisamment générique pour subsumer plusieurs cas de la base de cas.

Toutefois les solutions n'ont pas toujours de sens, notamment si les généralisations remonte jusqu'à $\top$. Aussi, avons nous considéré comme valables des appariements entre sommets ayant des concepts "relativement proches" dans la hiérarchie des concepts $\mathcal{H}_C$. Chacun des graphes $\texttt{org(source}_i\texttt{)}$ est apparié avec les parties de `org(cible)` correspondantes. Le chemin de similarité entre une source et cible est calculé sur la base de l'instanciation et des généralisations effectuées pour passer de `org(source)` à `idx(org(source))` et des

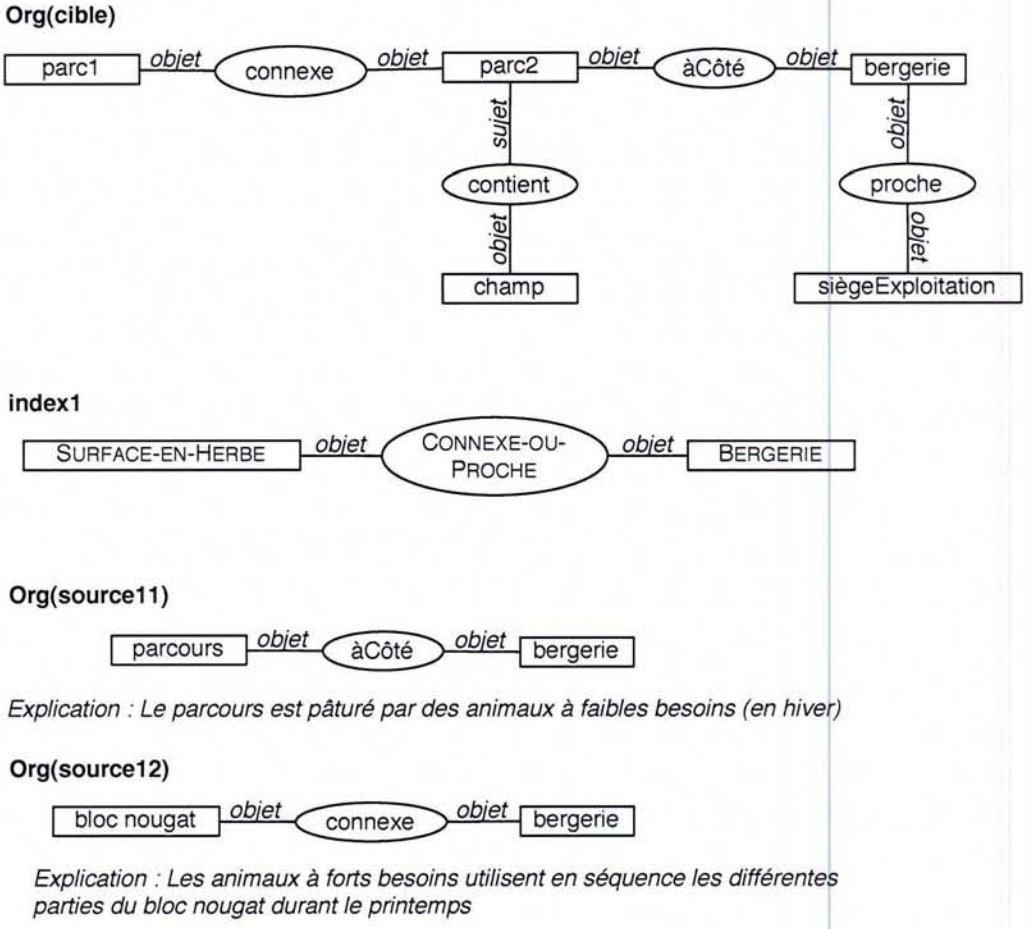


FIG. 5.5 – Exemple de remémoration à l'aide de la classification dure des graphes.

spécialisations suivies d'une instanciation menant de `idx(org(source))` à `org(cible)`. Un coût peut être associé à chacune des transformations par généralisation ou spécialisation sur chacun des sommets des graphes appariés afin de déterminer un coût associé au chemin de similarité permettant ainsi de classer les différents cas sources trouvés. Ce coût peut être calculé selon la formule de distance relative au sein d'une hiérarchie définie dans [37] :

$$dist(C_i, C'_i) = \frac{2\ profondeur(LCS_{(C_i,C'_i)})}{profondeur(C_i) + profondeur(C'_i)}$$

Où $dist(C_i, C'_i)$ est comprise entre 0 et 1. Elle vaut 1 lorsque le concept C_i associé au sommet S_i du graphe source est le même que le concept C'_i associé au sommet S'_i du graphe cible. Lorsque les sommets sont incomparables (le plus petit subsumant commun est $\top$), la distance vaut zéro. La distance pour un graphe représente la moyenne des distances entre concepts appariés :

$$\frac{\sum^{n} dist(C_i, C'_i)}{n}$$

5.4.4 Transformation des graphes

Le processus de classification dure est fondé sur un appariement exact entre la source et la cible. Cependant, ce processus ne donne pas toujours des résultats satisfaisants. Si l'on désire obtenir de nouveaux appariements pour proposer plus d'explications, il est nécessaire d'envisager de *transformer* le cas cible de façon à ce que l'opération de classification dure puisse être ré-appliquée pour donner des résultats complémentaires. En d'autre termes : $classification\ élastique = transformations + classification\ dure$.

Dans ce cas, le chemin de similarité est construit de la façon suivante :

$$\texttt{org(source)} \sqsubseteq \texttt{idx(org(source))} \sqsupseteq \texttt{T(org(cible))} \simeq \texttt{org(cible)}.$$

5.4.4.1 Règles d'inférence spatiale

Les transformations `T` qui peuvent être appliquées à `org(cible)` dans notre système ROSA correspondent à des règles d'inférence spatiale. D'une façon plus générale, les transformations possibles dépendent du domaine de l'application envisagée. Les règles d'inférence spatiale que nous manipulons peuvent être classées suivant trois catégories :

- Les règles de voisinage : une relation peut être transformée en une relation voisine. Les graphes de voisinage ont été décrits pour les relations de la théorie RCC-8 dans [24] comme présenté figure 5.6.

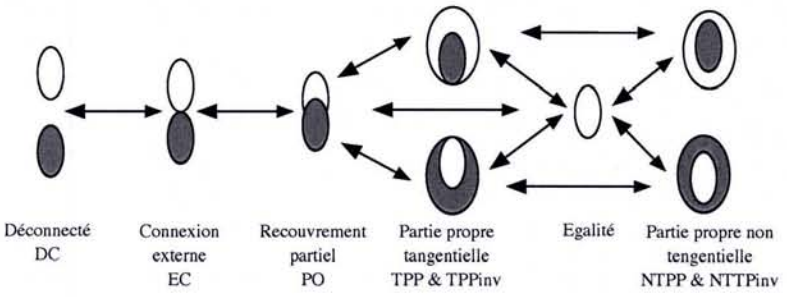


FIG. 5.6 – Graphe des relations de voisinage pour les relations de la théorie RCC-8 [24].

- Les règles de composition de relations : une relation entre deux régions peut être inférée à partir de la composition de relations entre ces deux régions et une troisième. Par exemple :
  - Ir-1 : *Si une région A est à l'intérieur d'une région B et que la région B est proche d'une région C alors la région A est proche de la région C.*

  Les compositions de relations de distance ont été étudiées dans [22] et les compositions de relations topologiques dans [33, 111].
- Les règles d'inférences spécifiques liées aux connaissances des experts du domaine. C'est notamment le cas pour des relations qui font référence à des propriétés spatio-fonctionnelles. Par exemple, nous pouvons citer les deux règles d'inférence spatiale suivantes illustrées par la figure 5.7 :
  - Ir-2 : *Si une surface B est comprise entre deux surfaces A et C, où A est connectée à B et B est connectée à C et si A et B peuvent être unifiées dans une unique surface A+B, alors A+B est connectée avec C.*

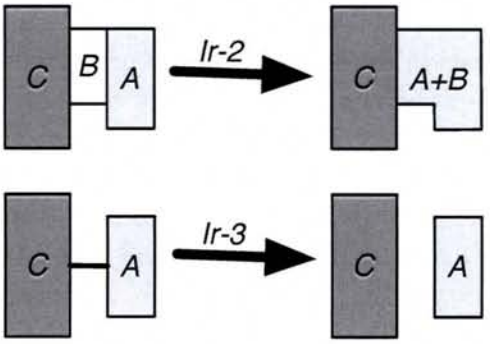


FIG. 5.7 – Deux exemples de règles de transformation spatiale.

- Ir-3 : *Si une ligne L relie une surface A à une autre surface C, et si L peut être supprimée, alors A et C sont déconnectés.*

Une fois appliquées les règles possibles, un nouveau cycle de remémoration peut être lancé sur les parties non expliquées du graphe cible ainsi transformé.

5.4.4.2 Utilisation des points de vue

Comme nous l'avons vu à la section 4.5.2, les cas peuvent traiter de problèmes avec des points de vue particuliers. Or la hiérarchie de connaissances du domaine décrit le point de vue particulier, et naturellement adopté par les agronomes, d'occupation du sol.

Par exemple, pour le point de vue DÉPACEMENTTROUPEAU, sont considérés comme similaires pour la fonction d'*attrait* les concepts suivants : ABRI, POINTD'EAU, CHAMP. En revanche pour le même point de vue, mais avec la fonction d'*obstacle* les concepts regroupés sont ROUTEPRINCIPALE et VILLE.

Ainsi, si le cas source remémoré traite de déplacement du troupeau avec un fonction d'attrait alors les concepts ABRI, POINTD'EAU et CHAMP seront considérés comme identiques i.e. $dist(C_i, C'_i) = 1$ entre ces concepts.

5.4.5 Classification élastique

Les deux opérations de base de la classification élastique sont les transformations et la classification dure. Les transformations sont appliquées de façon à modifier les relations et par là même la structure du graphe. Une transformation n'est appliquée que dans un contexte précis : type de relations impliquées, types des objets mis en relation, échelle de description considérée. L'application d'une transformation dans ces conditions doit garantir l'adaptabilité de la solution proposée.

Pour notre exemple présenté figure 5.5, certaines parties de `org(cible)` n'ont pas été appariées à des cas de la base de cas. On déclenche alors le processus de classification élastique en recherchant les règles d'inférence spatiale qui peuvent être appliquées sur des `s-relations` qui ne sont pas encore incluses dans une explication i.e. contient, connexe, loin. Pour le sommet `contient` la règle `Ir-1` peut être appliquée : le sommet `champ` est lié au sommet `parc2` par la relation CONTENIR et le sommet `parc2` est lié au sommet

bergerie par la relation TRÈS-PROCHE. Selon la règle d'inférence Ir-1, les sommets champ et bergerie peuvent être liés par une nouvelle relation PROCHE comme présenté figure 5.8.

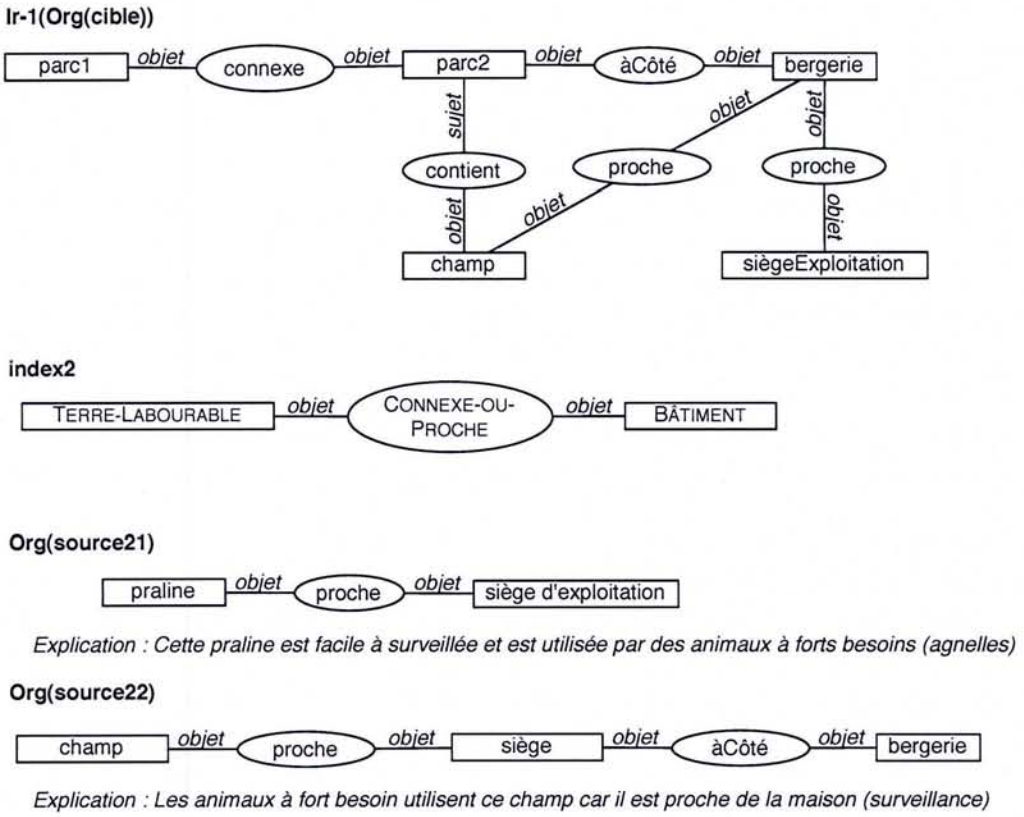


FIG. 5.8 – Exemple de remémoration à l'aide de la classification élastique des graphes.

Le graphe transformé Ir-1(cible) est le résultat de l'application de la règle d'inférence spatiale Ir-1 sur le Gos org(cible). Une fois ce nouveau Gos classé dans $\mathcal{H}_{idx}$, un nouveau graphe index index2 est apparié au sous-graphe (champ, proche, bergerie) de Ir-1(cible). Alors, ce dernier est comparé aux deux instances org(source21) et org(source22) de l'index index2 :

- De Ir-1(cible) à org(source21), la BERGERIE est généralisée en BÂTIMENT puis spécialisée en SIÈGE D'EXPLOITATION ; CHAMP est spécialisé en TERRE LABOURABLE puis spécialisé en PRALINE. Le coût du chemin de similarité est donné par les quatre opérations de généralisation / spécialisation plus le coût de l'application associée à la règle d'inférence ayant permis de modifier le graphe.
- De Ir(cible) à org(source22), seule la généralisation / spécialisation BERGERIE $\sqsupseteq$ BÂTIMENT et BÂTIMENT $\sqsubseteq$ SIÈGE est associée au coût de la transformation.

5.4.6 Pistes pour l'adaptation

Une fois que la sélection des différents cas sources, associés à différentes parties du graphe cible, est effectuée, le processus d'adaptation est activé. Les cas remémorés sont

alors classés en fonction de leur coût. Pour chaque graphe remémoré, on parcourt le chemin de similarité en réalisant des opérations similaires (généralisation, spécialisation, transformation) sur l'explication pour l'adapter. La première opération consiste à généraliser `exp(source)` pour obtenir une explication `exp(idx(source))`. Puis, on poursuit l'adaptation en fonction des opérations du chemin de similarité afin d'obtenir l'explication associée à l'organisation de cible `exp(cible)`.

De nombreux problèmes restent à résoudre pour fournir un processus d'adaptation efficace et générique à notre système ROSA. Comme nous l'avons expliqué, l'adaptation dans le RÀPC dépend principalement de l'appariement entre les organisations spatiales des sources $\texttt{org(source}_x\texttt{)}$ et de la cible `org(cible)` considérée. Dans notre système ROSA l'appariement des `gos` est réalisé par la classification de concepts mais l'application de règles d'inférences spatiales reste à résoudre. Nous avons formalisé une dizaine de règles immédiatement reconnaissable et quelques règles liés au domaine comme dans le cas de la présence d'un pré qui fait tampon. Les règles doivent jouer un rôle important dans l'élaboration des chemins de similarité et donc aussi dans le processus d'adaptation. Plus de règles devront être formalisées et testées pour évaluer les conditions de leurs applications. Ces règles devront être exprimées dans le but de permettre l'adaptation des explications d'une `exp(source)` vers une `exp(cible)`.

Actuellement, le système délivre un ou plusieurs chemins de similarité reliant les `Org(sources)` remémorés à une ou des parties de `org(cible)`. Les agronomes ont la charge d'accepter ou de refuser ces chemins et de construire les explications correspondantes. Une dizaine d'exploitations ont été étudiées sur les deux régions retenues (Lorraine et Causses) pour l'étude. Chaque exploitation à donné lieu à la description d'une dizaine de cas décrivant une organisation spatiale à laquelle est liée une explication de fonctionnement. L'étude plus poussée sur le causse Méjan à permis d'identifier une trentaine de cas différents enregistrés dans la base de cas.

Chapitre 6

Implantation de ROSA

Sommaire

6.1 Architecture du système ROSA 100
6.2 Modèle des données et persistance 100
6.2.1 La base terminologique 101
6.2.2 La base factuelle 102
6.3 ROSA-Interface 104
6.4 ROSA-RàPC 105
6.5 Conclusion 108

Après avoir représenté les connaissances et envisagé les stratégies de remémoration et d'adaptation, nous allons présenter le système ROSA tel que nous l'avons développé pour une utilisation par les agronomes. Nous décrivons le modèle des données, et comment l'interface utilisateur permet l'édition des connaissances. Puis, nous verrons quelles sont les limites de l'application des raisonnements de remémoration à l'aide des LD et comment nous avons développé une interface à base d'objets pour permettre la classification des graphes.

6.1 Architecture du système ROSA

Le système ROSA couvre deux fonctionnalités distinctes et complémentaires : une interface utilisateur pour les entrées et sorties d'informations et un système de manipulation des informations dans le cadre du RÀPC. La figure 6.1 présente l'architecture générale de notre application. Cette architecture est articulée autour des modules ROSA-Interface et ROSA-RÀPC :

- ROSA-Interface permet la saisie et l'édition des données sur les exploitations et la définition des connaissances du domaine. Chaque élément de connaissance manipulé par ROSA-Interface est instancié par une classe du modèle des données implanté dans ROSA-RÀPC. D'autre part, l'interface permet la visualisation des résultats des raisonnements mis en œuvre par ROSA-RÀPC et offre l'accès aux fonctions de RACER pour des besoins de test.
- ROSA-RÀPC communique avec le système RACER via l'API JAVA[22] de ce système. ROSA-RÀPC implante un modèle des données, à l'aide des classes JAVA qui est en correspondance avec le modèle des données de RACER : concepts, rôles, individus, relations entre individus. Ce modèle des données est relativement pauvre comparé à la structure des données décrites au chapitre précédent. Nous avons donc spécialisé la hiérarchie de classe JAVA pour en rendre compte. Toutes les données sont enregistrées au format XML dans un fichier qui est chargé à l'ouverture et qui permet d'instancier le modèle des données tant JAVA que RACER.

6.2 Modèle des données et persistance

Pour enregistrer les descriptions des concepts, nous avons opté pour le formalisme, fondé sur l'utilisation de XML, proposé par J. Euzenat : "*Description Logics Modeling Language*" (DLML[23]) [34]. Depuis, d'autres formats ont vu le jour et même RACER propose dans sa dernière version (1.7.7) de lire directement les descriptions à partir d'un fichier dans les formats suivants : OWL, DAML+OIL, RDFS ou RDF avec cependant quelques restrictions sur certaines expressions des langages OWL et DAML+OIL [46]. Au moment du choix du format de sauvegarde des données de l'application, ces autres formats (OWL,

[22] API : Application Programming Interface. L'interface de programmation d'application permet de définir les méthodes (JAVA) à utiliser pour avoir accès aux fonctionnalités du système pour lequel l'API est fourni.

[23] `http://co4.inrialpes.fr/xml/dlml/`

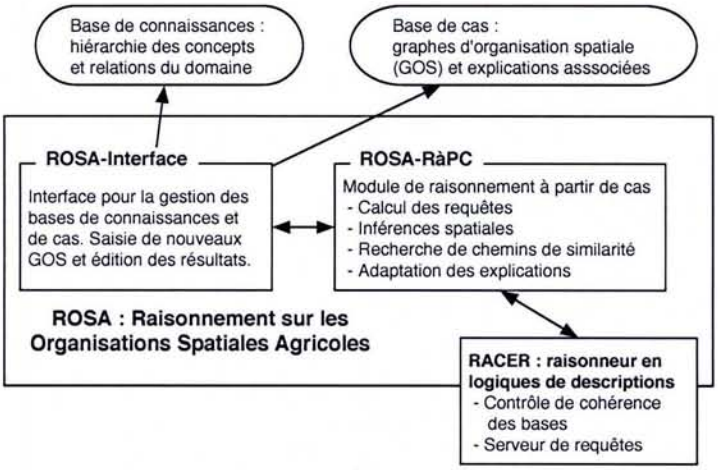


FIG. 6.1 – Architecture du système de Raisonnement sur les Organisations Spatiales Agricoles (ROSA).

DAML+OIL) n'avaient pas leur maturité actuelle et DLML proposait plusieurs feuilles de style XSLT [24] pour permettre la traduction d'une famille de logiques de descriptions vers une autre famille ou bien directement sous une forme exploitable par différents systèmes de LD. Maintenant, la conversion du format DLML au format DAML+OIL et inversement est proposée aux utilisateurs de DLML.

DLML propose une architecture de document fondée sur une balise racine notée `<dl:TERMINOLOGY>`. Cette balise englobe la description d'une base de connaissances décrite selon le formalisme des LD en introduisant la description des concepts, des rôles et des individus. Nous avons gardé cette structure pour la description des concepts et des rôles uniquement. Nous avons défini une nouvelle balise racine de notre document : `<dl:KNOWLEDGEBASE>`. Cette racine intègre deux branches distinctes : `<dl:TERMINOLOGY>` pour la représentation des éléments de la TBox et `<dl:ASSERTION>` pour les faits représentés dans la ABox.

6.2.1 La base terminologique

Quatre types d'objets peuvent être enregistrés dans une base terminologique. Il s'agit des concepts et des rôles qui peuvent être primitifs ou définis. Dans le cadre de l'enregistrement au format DLML, les premiers concepts primitifs définis dans la hiérarchie $\mathcal{H}_\mathcal{C}$ sont enregistrés avec la succession de balises suivantes :

```
<dl:CPRIM>
<dl:CATOM> SOMMET </dl:CATOM>
<dl:ANYTHING/>
</dl:CPRIM>
```

La balise `<dl:CPRIM>` définit la construction d'un concept primitif. Le concept représenté

[24] Langage de mise en forme et de traduction pour XML (stylesheet language for XML Transformations) `http://www.w3.org/TR/xslt`

a pour nom SOMMET et est déclaré comme étant un concept atomique, ici simplement subsumé par le concept $\top$ représenté par la balise `<dl:ANYTHING>`.

Un concept défini, comme par exemple, le concept RELATION se représente comme suit :

```
<dl:CDEF>
  <dl:CATOM>RELATION</dl:CATOM>
  <dl:AND>
    <dl:CATOM>SOMMET</dl:CATOM>
    <dl:ALL>
      <dl:RATOM>arc</dl:RATOM>
      <dl:CATOM>RÉGION</dl:CATOM>
    </dl:ALL>
  </dl:AND>
</dl:CPRIM>
```

La balise `<dl:CDEF>` déclare la création d'un concept défini. Ce concept a pour nom RELATION. Il est dérivé du concept SOMMET et doit posséder un *arc* dont le co-domaine est le concept RÉGION.

Deux exemples de description de rôles sont présentés ci-dessous. Pour le premier rôle, la balise `<dl:RDEF>` déclare la construction d'un rôle défini. Le nom du rôle est déclaré par la balise `<dl:RATOM>`. Ensuite, les propriétés associées à un *arc* sont précisées : un domaine (`<dl:DOMAIN>`) définit l'origine de l'*arc* pris parmi les individus de type RELATION et un co-domaine (`<dl:RANGE>`) définit sa destination, prise dans l'ensemble des individus de type RÉGION. La deuxième déclaration construit le rôle *objet* comme étant dérivé du rôle *arc* à l'aide de la balise `<dl:PARENT>`.

```
<dl:RDEF>
  <dl:RATOM>arc</dl:RATOM>
  <dl:DOMAIN>Relation</dl:DOMAIN>
  <dl:RANGE>Région</dl:RANGE>
</dl:RDEF>

<dl:RPRIM>
  <dl:RATOM>objet</dl:RATOM>
  <dl:PARENT>arc</dl:PARENT>
</dl:RPRIM>
```

6.2.2 La base factuelle

Nous aurions pu utiliser le formalisme proposé par DLML pour l'enregistrement des connaissances sur les exploitations (notamment leur `gos`) sous forme d'une liste d'individus à l'aide des balises :

```
<dl:INDIVIDUAL name="nomIndividu">
  <dl:CATOM>nomConceptAtomique</dl:CATOM>
</dl:INDIVIDUAL>
```

Les relations entre individus ne sont pas proposées actuellement dans la définition de DLML mais pourraient avoir la forme suivante :

```
<dl:RASSERTION>
  <dl:CATOM>nomInstanceOrigine</dl:CATOM>
  <dl:CATOM>nomInstanceDestination</dl:CATOM>
  <dl:RATOM>nomRôleAtomique</dl:RATOM>
<dl:RASSERTION>
```

Cependant, nous avons opté pour une structuration plus forte et par là même plus explicite pour le format XML d'enregistrement des connaissances sur les individus. Nous laissons donc le soin à l'application de gérer l'ensemble des individus et de leurs relations pour les représenter dans le système de LD.

Par abus de langage, nous avons appelé chorème l'instance de GOS décrivant une exploitation. Cette balise `<CHOREME>` possède les attributs suivants :

- nom : pour définir le nom de l'instance représentant l'exploitation enregistrée,
- concept : il permet de typer l'instance de graphe avec le concept GOS,
- image : définit un lien sur le dessin du chorème dont le graphe est une traduction,
- commentaire : pour permettre au modélisateur agronome d'enregistrer les commentaires ou remarques qu'il juge utile d'apporter.

Cette première balise contient la liste des sommets du graphe à travers l'utilisation de balises `<SENTITE>` et `<SRELATION>` qui possèdent les mêmes attributs que la balise `<CHOREME>` à l'exception de l'attribut Image. Les balises `<SRELATION>` contiennent à leur tour des balises `<ARC>` permettant de spécifier à l'aide de deux attributs les `s-entités` mises en relation en précisant leur rôle dans la relation.

Au chapitre précédent, nous avons présenté un exemple (tableau 5.3) ne décrivant qu'une très petite partie du graphe de l'organisation spatiale d'une exploitation. Pourtant déjà, nous avons constaté la faible lisibilité de cette syntaxe et le nombre élevé de lignes à écrire notamment pour déclarer les instances appartenant au graphe. La formulation proposée ici, en plus d'augmenter la lisibilité, permet de déplacer la gestion des relations vers un niveau applicatif, déchargeant ainsi l'utilisateur d'une tâche laborieuse d'identification des relations entre les individus. Le tableau 6.1 présente l'enregistrement fait par l'application au format XML de l'exemple décrit dans le tableau 5.3.

Un cas est défini pour une exploitation donnée. Sa description est donc naturellement ajoutée au sein de la balise `<CHOREME>` de description des exploitations. L'exemple du tableau 6.2 est un cas associé à la description de l'exploitation présentée au tableau 6.1. Le cas est un `gos-e` défini par une liste des sommets impliqués dans l'organisation spatiale considérée et une explication. Seuls les sommets `s-relations` sont enregistrés ici. Les autres sommets sont implicitement enregistrés avec les arcs sortant de chacun des sommets relation du `gos-e`.

L'explication donnée sous forme de simple phrase, ne peut pas être utilisée telle quelle. Aussi, nous proposons de faire décrire l'explication par l'utilisateur sous la forme d'une liste d'unités linguistiques qui constituent la phrase explicative du cas. Chaque unité faisant référence à un des sommets du `gos` y est associée. Les autres unités linguistiques sont simplement énoncées sans référence à un sommet. L'exemple du tableau 6.2 présente le cas comme décrivant une organisation spatiale, à l'échelle du territoire de l'exploitation,

```
<dl:ASSERTION>
  <CHOREME concept="Choreme" nom="exploit1" image="ressources/Exploit1.jpg">
    <SENTITE concept="Siège d'exploitation" nom="siègeExploit1"/>
    <SENTITE concept="Bergerie" nom="bergerieExploit1"/>
    <SENTITE concept="Champ" nom="champClos1">
      <PROPRIETE concept="Fort potentiel"/>
      <PROPRIETE concept="Clôturé"/>
    </SENTITE>
    <SRELATION concept="Etre à Côté" nom="àCôté1">
      <ARC entite="siègeExploit1" role="objet"/>
      <ARC entite="champClos1" role="objet"/>
    </SRELATION>
    <SRELATION concept="Etre à Côté" nom="àCôté2">
      <ARC entite="siègeExploit1" role="objet"/>
      <ARC entite="bergerieExploit1" role="objet"/>
    </SRELATION>
  </CHOREME>
</dl:ASSERTION>
```

TAB. 6.1 – Traduction, au format XML de notre application, de l'exemple présenté dans le tableau 5.3.

fondée sur deux relations reliant trois `s-entités`, avec une explication constituée de huit unités linguistiques dont cinq sont en relation avec les sommets du `gos-e`.

6.3 ROSA-Interface

Pour saisir ou modifier les connaissances, il nous faut une interface qui réponde aux besoins de représentation tels que nous venons de les décrire ainsi qu'aux besoins des futurs utilisateurs qui sont des agronomes. Ce dernier point est une vaste question que nous n'abordons pas ici.

Il existe plusieurs applications génériques permettant la manipulation des connaissances. Nous pouvons citer OIL-ed[25], un éditeur graphique de connaissances terminologiques qui fait suite aux travaux de formalisation du langage de représentation DAML+OIL. PROTÉGÉ-2000 [26] semble être le projet le plus abouti actuellement. Il a pour but de fournir une interface utilisateur de haute qualité pour la manipulation de concepts et d'individus et a maintenant pris une importance considérable. En effet, PROTÉGÉ-2000 permet la gestion de bases de connaissances tant terminologiques que factuelles et offre actuellement une interface avec des raisonneurs en LD pour les deux systèmes actuellement en développement et en concurrence que sont FACT et RACER. De plus, PROTÉGÉ-2000 est une plate-forme construite pour permettre l'ajout facile de fonctionnalités nouvelles par adjonction de plug-in.

[25] `http://oiled.man.ac.uk/`
[26] `http://protege.stanford.edu/`

```
<CAS explication="L'agriculteur utilise ce champ à fort potentiel (ressource),
    clôturé (sans garde) proche du siège (surveillance) et proche de la bergerie
    (déplacements limités) pour y mettre les animaux à très forts besoins."
    nom="Cas Puel 1" Echelle="Echelle Territoire">
  <NOEUD nom="àCôté1"/>
  <NOEUD nom="àCôté2"/>
  <EXPLICATION>
    <UL texte="L'agriculteur utilise" sommet=""/>
    <UL texte="ce champ à fort potentiel (ressource), clôturé (sans garde)"
      sommet="champClos1"/>
    <UL texte="proche" sommet="àCôté1"/>
    <UL texte="du siège (surveillance)" sommet="siègeExploit1"/>
    <UL texte="et" sommet=""/>
    <UL texte="proche" sommet="àCôté2"/>
    <UL texte="de la bergerie (déplacements limités)" sommet="bergerieExploit1"/>
    <UL texte="pour y mettre les animaux à très forts besoins." sommet=""/>
  </EXPLICATION>
</CAS>
```

TAB. 6.2 – Exemple de description d'un cas fondé sur le graphe du tableau 6.1.

PROTÉGÉ-2000 semble aujourd'hui être un bon compromis pour l'aspect gestion de bases de connaissances dans le formalisme des LD. Cependant, nous n'envisageons pas son utilisation dans le cadre de notre système. En effet, nous voulons une application accessible à des utilisateurs non formés au formalisme des LD, au moins pour la saisie des connaissances factuelles. En revanche, son utilisation devrait être envisagée dans un futur système pour la gestion des bases de connaissances terminologiques, ce d'autant plus que PROTÉGÉ-2000 fonctionne avec les systèmes de LD actuels et permet des sauvegardes selon différents formalismes dont les langages DAML+OIL et OWL.

Nous avons donc entrepris de développer notre propre application pour saisir les connaissances sur les exploitations. ROSA-Interface permet de visualiser, de modifier et de saisir toutes les informations nécessaires. L'application se compose d'une fenêtre unique (figure 6.2) offrant plusieurs vues accessibles à tout moment à travers la sélection des différents onglets. Cette application permet de créer, ou de charger à partir d'un fichier, une base de connaissances. Elle permet aussi de se connecter au serveur de LD RACER pour y charger les connaissances et assurer l'accès à ses fonctionnalités. L'annexe B donne une description détaillée de l'interface de ROSA et de son utilisation.

6.4 ROSA-RàPC

ROSA-RÀPC assure l'interface entre les connaissances saisies via ROSA-Interface et le système de LD RACER et gère les étapes du cycle du raisonnement à partir de cas. Comme nous l'avons vu au chapitre précédent, la remémoration dans ROSA est fondée sur la classification de graphes. Cette classification est assurée par une fonction d'appariement entre

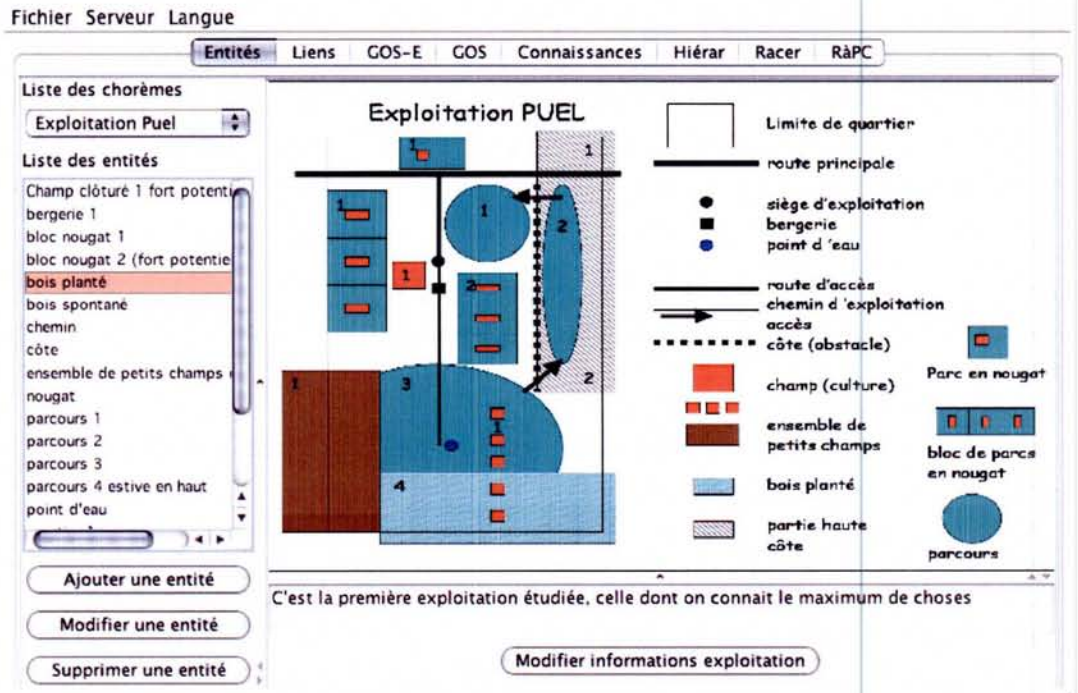


FIG. 6.2 – Un exemple de saisie ou de modification des entités d'un chorème dans ROSA-Interface.

graphes. Les graphes manipulés sont des graphes orientés conformément à la relation entre instances définie par les rôles des LD. Les sommets des graphes sont des instances munies d'un identifiant unique et d'un type défini par un concept de la hiérarchie $\mathcal{H}_{\mathcal{CE}}$ pour les sommets `s-entités` ou de la hiérarchie $\mathcal{H}_{\mathcal{CR}}$ pour les sommets `s-relations`. La structure des graphes est définie à l'aide des rôles particuliers *objet* et/ou *sujet* qui définissent le statut des `s-entités` dans les relations. Chaque type d'objet dans ROSA – concept, rôle, chorème, s-entité, s-relation, arc – est manipulé par une classe JAVA correspondante. Les différents objets sont instanciés lors de la lecture des bases de connaissances depuis le fichier XML, ou lors de leur création via ROSA-Interface. ROSA-RÀPC permet alors le chargement de la description des concepts, des rôles, des instances et de leurs liens dans le système RACER.

Il est facile en logiques de descriptions de trouver un sommet S de G plus spécifique qu'un sommet S' de H. Il est aussi possible de trouver un sommet S_1 de G tel que S_1 soit lié à S_2 via un rôle r_1, et qui soit plus spécifique qu'un sommet S'_1 lié à à un sommet S'_2 de H. De même, il est facile de déterminer une relation R de G reliant deux sommets (relation binaire avec S_1 et S_2), trois sommets (relation ternaire avec S_1, S_2 et S_3) ou plus, telle que son image dans H soit plus spécifique tout en conservant la structure de la relation. En revanche, si S_1 est connecté à des sommets S_2 et S_3 via deux relations R_1 et R_2, il n'est pas possible de garantir à l'aide des LD que sa projection soit bien de la forme S'_1 lié à S'_2 par une relation R'_1 et à S'_3 par une relation R'_2 car l'opérateur "*Same-As*" entre instances n'existe pas. On peut très bien se retrouver avec quelque chose de la forme

S'_1 lié à S'_2 par une relation R'_1 et S'_3 lié à S'_4 par une relation R'_2 comme présenté figure 6.3. Rien ne garantit que S'_1 et S'_4 soient un seul et même sommet. La raison de cette difficulté tient au fait que les LD sont définies dans un cadre ensembliste où les individus sont classés en fonction des propriétés qu'ils partagent mais ne sont pas accessibles en fonction de leur position les uns par rapport aux autres.

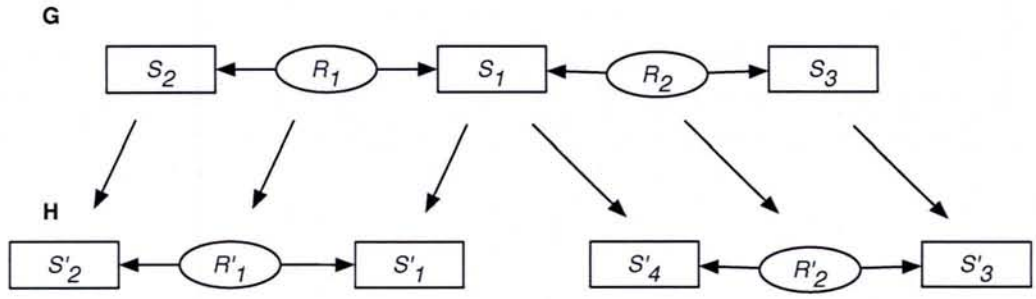


FIG. 6.3 – Difficulté d'appariement de graphe dans le cadre des LD ; les sommets S'_1 et S'_4 doivent représenter un seul et même individu, ce que ne permet pas de garantir le système de LD.

Aussi avons nous proposé d'assurer le contrôle de la structure des graphes à l'aide d'une fonction écrite en JAVA en nous appuyant sur le raisonneur RACER pour retrouver les sommets et les arcs ayant les propriétés désirées. La fonction d'appariement de deux graphes H et G se déroule en deux étapes :

1. Pour chacun des sommets (et des arcs) de H, une liste des sommets (et des arcs) appariables de G est constituée. Ces listes sont déterminées en fonction du type de similarité considérée en s'appuyant sur les capacités de classification offertes par le système RACER.
2. Par un parcours en profondeur de H, on reconstitue sa structure à l'aide des listes de sommets appariables. Au terme du parcours, les différents sous-graphes de G isomorphes à H sont retournés.

Au terme du parcours du graphe H, nous avons une liste de sous-graphes de G appariés à H. Pour chaque graphe, chaque appariement des sommets sur H et G peut être visualisé avec le chemin parcouru pour passer du concept d'un sommet de H au concept d'un sommet de G.

Deux types de classification ont été implantés. La première assure une classification dure. Le concept de chaque sommet de H est plus général que le concept associé au sommet apparié de G. Le deuxième type correspond à la classification "semi-dure" pour laquelle on recherche le concept, dans la hiérarchie $\mathcal{H}_\mathcal{C}$, le plus spécifique subsumant les concepts des sommets appariés de H et G.

Les graphes réponses sont alors classés dans une liste en fonction de la similarité calculée selon le principe développé à la section 5.4.3. La similarité pour un graphe est donnée par la valeur moyenne des distances sur chaque sommet. Cette valeur est comprise entre 0 (graphe non comparable) et 1 (graphe identique).

L'introduction des points de vue permet d'étendre les listes de sommets appariables en fonction des concepts considérés alors comme comparables. La similarité dans ce cas est considérée comme égale à 1, i.e. le coût n'est plus calculé en fonction de la position respective des concepts dans la hiérarchie $\mathcal{H}_C$.

6.5 Conclusion

Le système que nous avons développé permet aux agronomes de saisir des graphes de description des exploitations et permet de saisir les cas de la base de cas. Ce système permet aussi d'éditer les hiérarchies de concepts et de rôles ainsi que les points de vue.

De plus, l'algorithme d'appariement entre graphes permet de classer les graphes. Deux types de classification ont été implantés et testés : la classification dure et la classification "semi-dure". Les résultats sont discutés au chapitre suivant.

Chapitre 7

Intérêts et usages pour les agronomes

Sommaire

7.1 L'évolution des chorèmes 110
7.1.1 Modification d'un chorème d'exploitation agricole 111
7.1.2 Le graphe pour expliquer les relations spatiales 113
7.1.3 Le graphe pour valider la représentation chorématique . . . 114
7.2 L'utilisation du système ROSA 115
7.2.1 La saisie d'un gos 115
7.2.2 La saisie d'un cas 116
7.2.3 Le RàPC dans ROSA 116
7.3 Usages du système 121

Au cours du travail d'acquisition des connaissances (présenté au chapitre 4), nous avons cherché à construire le modèle des connaissances et des raisonnements à implanter dans ROSA. Le travail d'acquisition de connaissances, par la formalisation des organisations spatiales à l'aide de graphes, a eu pour conséquence une amélioration, une formalisation sensible de l'outil graphique des agronomes.

L'utilisation du système a permis de saisir les connaissances sur les premières exploitations étudiées. Les connaissances enregistrées correspondent aux `gos` des exploitations ainsi que les premiers cas de la base de cas. Enfin, nous avons pu tester les différents modes de remémoration implantés pour proposer des explications sur une nouvelle exploitation. Cette utilisation de l'outil ROSA a mis en évidence des faiblesses et lacunes qui devront être prises en compte au cours des développements futurs de notre outil.

7.1 L'évolution des chorèmes

La phase d'apprentissage, d'acquisition de connaissances est une phase importante, et jamais totalement aboutie, dans le développement d'un système à bases de connaissances. La proposition de transformer les chorèmes en graphes a pris naissance et s'est développée au fur et à mesure. La modélisation des connaissances (et des cas) à représenter dans le système s'est déroulée durant des séances de travail entre agronomes et informaticiens suivies de moments de maturation, de validation des modélisations proposées et de résolution des problèmes soulevés par certaines propositions ou constatations. Au cours de ces séances, les outils mis en jeu sont d'une part les chorèmes - représentations construites par les agronomes pour schématiser l'organisation spatiale et le fonctionnement des exploitations agricoles - d'autre part, les graphes utilisés par les informaticiens pour modéliser les organisations spatiales. Le passage de l'un à l'autre est un moyen de mettre à jour et de formaliser les connaissances des agronomes à propos des exploitations étudiées. Ces séances de travail se sont tenues comme un entretien ouvert s'appuyant sur les représentations graphiques [62].

Au cours des séances de travail en commun, trois phases dans l'acquisition des connaissances ont été observées :

- La constitution de graphes traduisant l'organisation spatiale des exploitations étudiées.
- Une analyse des organisations spatiales représentées par le schéma offre l'occasion de déterminer les explications de fonctionnement qui y sont associées. L'association d'une explication à une structure spatiale permet de définir un cas. La partie problème du cas est représentée par le graphe décrivant l'organisation spatiale. La solution correspond à l'explication fournie.
- L'analyse des dénominateurs communs entre plusieurs chorèmes d'exploitation permet de comparer des organisations spatiales et de déterminer celles qui sont similaires et qui peuvent donner lieu à des explications semblables au prix de certaines adaptations.

Ces phases s'organisent dans des cycles itératifs d'interrogation et de modification des chorèmes et des graphes correspondants, pour aboutir à une convergence, un accord, entre agronomes et informaticiens.

7.1.1 Modification d'un chorème d'exploitation agricole

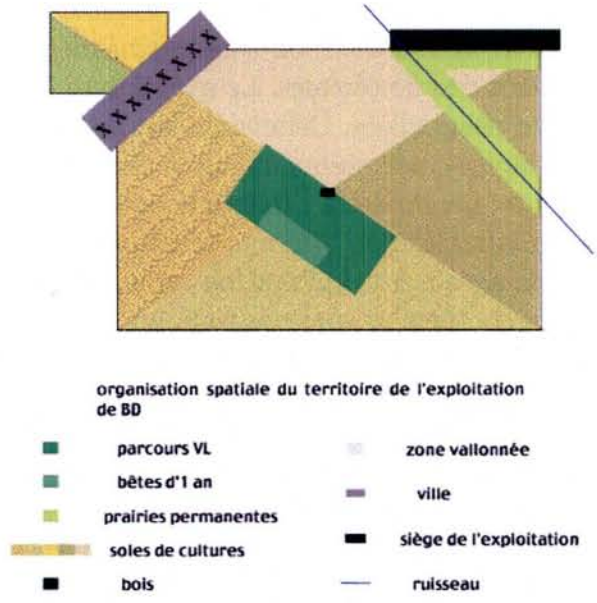


FIG. 7.1 – Chorème représentant l'organisation spatiale d'une exploitation agricole de l'est de la France.

Le chorème présenté figure 7.1 met en évidence une exploitation compacte de polyculture élevage en Lorraine. Deux types de surfaces sont distingués et recouvrent deux activités différentes de l'exploitation. D'une part, les surfaces toujours en herbe, centrées autour du siège de l'exploitation, sont destinées à la pâture des animaux. La principale activité de l'exploitation est assurée par la production de lait. D'autre part, il y a des surfaces de culture de céréales destinées en partie à la vente, l'autre partie assurant le complément alimentaire des animaux. La sole cultivée est partagée en quatre secteurs, correspondant à une rotation du système de culture. L'existence d'une ville et d'un ruisseau est mentionnée. Ce chorème d'exploitation est synthétique, il rend compte des grandes logiques de fonctionnement au niveau de l'exploitation.

Le schéma chorématique est construit pour rendre compte de l'organisation spatiale et du fonctionnement de l'exploitation agricole modélisée. Pour comprendre cette représentation et la traduire sous forme de graphe, l'informaticien "joue" le rôle du naïf. Sa position est de dire "Je ne comprends pas" afin d'obtenir une reformulation dans des termes qui mettent à jour les connaissances sous-jacentes. De même, un autre agronome, ayant une culture différente, oblige le créateur du chorème à expliciter son contenu. Le travail d'explicitation permet de formuler dans le détail les connaissances mobilisées et les raisonnements ayant abouti au dessin. Les ambiguïtés sont levées.

Ce travail permet de mettre en lumière, de réduire, de lever les imprécisions des chorèmes d'exploitations. Ces imprécisions sont principalement dues au fait que le dessinateur doit faire des choix quand il produit un chorème, afin de rendre le dessin synthétique et lisible. Ces choix sont propres à la personnalité du dessinateur, à sa façon de traduire en graphique sa compréhension du fonctionnement de l'exploitation. De plus, les chorèmes

contiennent beaucoup d'implicites liés à la région dans laquelle est située l'exploitation représentée, et une difficulté majeure tient à la variété des sens d'un même terme (par exemple, "champ") d'une région à l'autre, voire d'une exploitation à l'autre.

Par exemple, dans le cas de la figure 7.1, on observe quatre secteurs autour du siège de l'exploitation destinés aux cultures de céréales. La question se pose de savoir si ces quatre secteurs ont une signification particulière. Définissent-ils quatre champs, ou bien plusieurs champs utilisés de façon différente, par exemple dans le cadre d'une rotation des cultures ? Les petites surfaces de culture comprises à l'intérieur des prairies permanentes sont-elles des surfaces de culture à individualiser ou bien une conséquence de la technique du dessin ? Enfin, la question se pose de savoir si l'exploitation est sur un territoire rectangulaire ou non ? De fait, les objets d'un chorème ne sont pas toujours facilement identifiables, à cause de schémas peu explicites ou de configurations ambiguës.

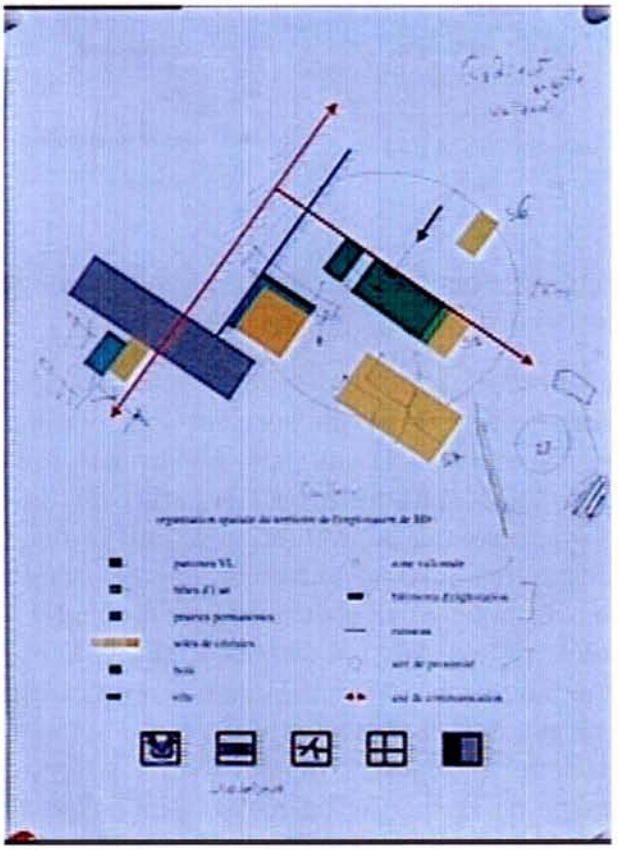


FIG. 7.2 – Nouveau chorème représentant la même exploitation que celle modélisée figure 7.1.

Ce travail conduit donc les agronomes à établir des règles supplémentaires pour uniformiser les règles de représentation chorématique. Par exemple, ils décident que la taille et la forme des parcelles, si l'information est significative, sont directement représentées dans le chorème. Sinon, les parcelles sont représentées comme des objets géométriques de même taille et de même forme. Les agronomes s'intéressent au territoire d'une exploitation et à ses relations avec son environnement. Le territoire de l'exploitation étant, en règle générale, constitué de parcelles dispersées sur une ou plusieurs communes, il n'y a aucune

raison d'avoir un dessin inscrit dans un espace fermé dont les éléments sont tous connexes comme dans la représentation initiale de la figure 7.1. Le résultat de l'application de ces nouvelles règles est présenté dans la composition chorématique re-dessinée figure 7.2. Ce chorème d'exploitation est plus détaillé. La topologie des parcelles et l'orientation sont plus respectés. La proportion des surfaces entre pâture et culture est plus conforme et la discontinuité des parcelles est rendue. La distance est plus explicite, d'une part par le figuré de l'aire de proximité, d'autre part par l'éloignement des parcelles de l'autre côté de la ville. Ce chorème d'exploitation est plus décomposé, il rend compte des modalités de fonctionnement au niveau des parcelles de l'exploitation. Les discussions autour de ce nouveau schéma chorématique ont abouti à l'élaboration des nouvelles règles plus fines suivantes :

- Chaque parcelle qui a une fonctionnalité spécifique pour l'agriculteur doit être représentée dans le chorème. C'est par exemple la parcelle réservée aux génisses d'un an ou le parc des vaches laitières.
- Si un groupe de parcelles répond aux mêmes fonctionnalités, par exemple de culture de céréales pour la vente, on symbolise dans le chorème un groupe de parcelles et on définit dans le graphe une entité "bloc de parcelles" ou "bloc de parcelles avec des céréales" : on gommera ainsi la position relative des parcelles les unes par rapport aux autres, en gardant l'indication qu'il existe plusieurs parcelles [17].

C'est ainsi que nous sommes arrivés à la composition chorématique finale présentée figure 7.3. Ce chorème est peu différent du précédent, il est juste épuré pour rendre compte des principales organisations : un bloc de prairies permanentes avec clôtures fixes et mobiles pour le pâturage des vaches laitières, à proximité du bâtiment d'exploitation, un bloc de parcelles en céréales correspondant à un même système de culture, un ensemble bois/rivière/prairie/culture ayant un mode de gestion spécifique (zone tampon) et des parcelles éloignées de l'autre côté de la ville qui fait obstacle.

La traduction des chorèmes en graphe oblige à expliquer les relations qui sont importantes et fait émerger les entités spatiales de gestion qui ont du sens pour l'exploitant. Cette démarche illustre l'intérêt de la combinaison d'outils et de la complémentarité de disciplines pour modéliser des phénomènes complexes. Elle ne s'appuie pas uniquement sur l'interaction entre des chercheurs de disciplines différentes ou sur la mobilisation conjointe d'outils. Elle repose sur une association continue entre des points de vue sur un territoire (les règles de fonctionnement des exploitations en sont l'expression) et des propriétés de l'espace (révélées par la structure de l'exploitation) pour aboutir à un modèle qui intègre le spatial et le fonctionnel, et qui de ce fait renvoie aux stratégies des exploitants agricoles, primordiales dans la transformation des territoires ruraux.

7.1.2 Le graphe pour expliquer les relations spatiales

Le choix des relations est parfois délicat. En particulier, le choix des relations de proximité ou d'éloignement est d'une grande importance. En effet, une surface peut être considérée comme éloignée si elle se situe à plus d'un kilomètre de l'exploitation si les animaux doivent s'y déplacer, alors qu'elle sera considérée comme proche si c'est une surface sur laquelle n'interviennent que des travaux motorisés. Ou encore, une surface peut être considérée comme éloignée, même si elle est physiquement proche, si un obstacle

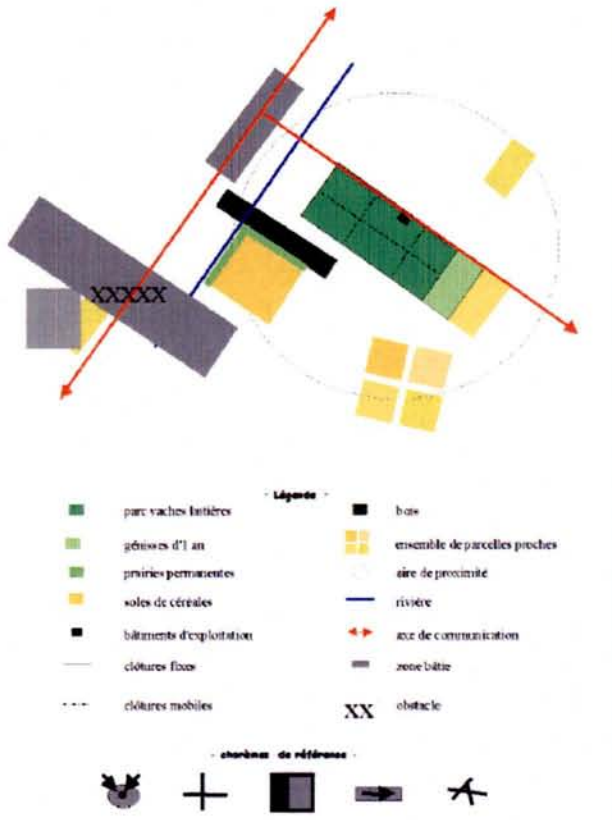


FIG. 7.3 – Composition chorématique finale de l'exploitation présentée figure 7.1 après la phase d'acquisition.

au déplacement existe pour l'atteindre (exemple d'une surface de culture de l'autre côté d'une agglomération par rapport au siège de l'exploitation, ou encore un pré situé de l'autre côté d'une route à fort trafic).

De même pour ce qui est du choix des relations pour décrire les relations existantes entre les éléments connexes de l'ensemble bois/rivière/prairie/culture au centre du chorème. Le choix exprimé à la section 4.4.2 montre bien que les relations considérées ne sont pas seulement topologiques, mais qu'elles ont du sens pour les thématiciens (ici les agronomes).

Ce travail sur les relations dans les chorèmes, et sur les graphes a permis d'enrichir la hiérarchie des relations spatiales qui ne se limite plus à de simples relations topologiques mais intègre des relations d'ordre spatio-fonctionnelles chargées de sens.

7.1.3 Le graphe pour valider la représentation chorématique

A l'issue du travail commun, les agronomes ont défini de nouveaux symboles pour rendre compte de situations spécifiques aux exploitations agricoles. Ces nouveaux éléments chorématiques donnent lieu à la définition de concepts particuliers recouvrant une réalité complexe. Au-delà du jeu de chorèmes élémentaires définis par les géographes, on construit ainsi un jeu de chorèmes "agronomiques" qui s'appliquent à la description des territoires

d'exploitation.

C'est le cas par exemple de la définition des blocs ou ensemble de parcelles certes séparées mais considérées comme un tout. C'est encore le cas lorsque l'on considère le choix fait entre des représentations de relations entre des chorèmes élémentaires ou bien une définition de nouveaux chorèmes complexes, associés à des concepts particuliers, tel que nous l'avons discuté à la section 4.4.3 pour les notions de NOUGAT et de PRALINE.

Par rapport aux chorèmes des géographes qui servent à décrire des régions entières, nous avons introduit une discontinuité pour représenter des interactions au niveau des exploitations. La représentation d'une discontinuité donne du sens si elle a une signification fonctionnelle (et non pas seulement structurelle).

Parfois l'on décompose des structures connexes, comme par exemple lorsque l'on considère la relation entre la prairie qui fait tampon entre une céréale d'une part et une rivière ou un bois d'autre part. D'autres fois on recompose un ensemble de parcelles de culture, bien que déconnectées, en un bloc, car ce qui compte, c'est la fonctionnalité remplie.

Les chorèmes d'exploitation sont dessinés en fonction du sens que l'on veut représenter. Et c'est le passage par le graphe qui a obligé les agronomes à expliciter leur représentation graphique. Notre travail a permis d'adapter une méthode des géographes, qui l'appliquent sur des régions pour la transformer en une méthode d'agronomes qui l'appliquent sur des territoires d'exploitations. L'échelle et la thématique ont été changées au profit d'une méthode générique qui rend compte de modèles spécifiques : chaque exploitation est unique, bien que la construction des chorèmes suive la même procédure.

7.2 L'utilisation du système ROSA

Au cours de plusieurs expériences, les agronomes ont pu utiliser le système. Une première expérience a permis de saisir des graphes d'organisations spatiales pour diverses exploitations avec la définition des premiers cas de la base de cas. Nous avons principalement mobilisé des connaissances sur les exploitations du causse Méjan pour définir les cas. Une autre expérience a permis de tester la remémoration et les problèmes liés aux différents modes de raisonnements implantés sur des exploitations du causse de Sauveterre.

7.2.1 La saisie d'un gos

La saisie d'un `gos` se fait normalement sur la base d'un chorème d'exploitation analysé par l'expert agronome chargé de l'enregistrer. Pour garder une trace du dessin, celui-ci peut être digitalisé et affiché en regard des informations rentrées dans le système.

Les graphes sont saisis en deux étapes. La première consiste à identifier l'ensemble des entités constitutives du chorème tel que décrit dans l'annexe B.2.1. Chaque entité spatiale correspond à un sommet `s-entité` de type RÉGION : un nom est affecté à l'instance et son type est choisi dans la hiérarchie des concepts d'occupation du sol (figure B.4). Des propriétés peuvent être ajoutées pour préciser les caractéristiques et par là même, le

type de l'instance. Par exemple, pour déclarer qu'un parcours a été clôturé, on définit le `parcours` avec le type PARCOURS et la propriété CLÔTURÉ.

La seconde étape consiste en la création du graphe proprement dit en identifiant les relations entre les entités. De la même façon que pour les entités, les relations sont réifiées et typées. Des propriétés peuvent aussi être ajoutées pour en préciser les caractéristiques. Les arcs sont ajoutés en sélectionnant les `s-entités` mises en relation. Ils sont étiquetés avec le rôle du `s-entité` dans la relation.

Cette façon de saisir est quelque peu laborieuse. Cependant la saisie des graphes par les agronomes eux même nous a révélé une volonté de donner plus de détails (ajout d'attributs notamment). D'autre part, nous avons observé un début d'auto-correction, dans la mesure où l'expert agronome se rend compte des imprécisions de son discours pouvant amener à des incohérences. L'envie de préciser, de détailler les situations apparaît très nettement.

7.2.2 La saisie d'un cas

L'analyse des chorèmes par les agronomes les mène à identifier des pratiques d'utilisation du territoire qui sont mises en relation avec les structures spatiales observées. Ainsi, des explications de fonctionnement sont associées à des parties de `gos` d'exploitation. Un sous-graphe et son explication associée définissent un cas constitué d'une partie problème : le sous-graphe et d'une partie solution : l'explication.

Pour saisir un cas, il faut donc sélectionner les sommets `s-relations` constitutives du sous-graphe considéré (et éventuellement des `s-entités` non connectées aux relations sélectionnées) et enregistrer l'explication dans la partie solution du cas comme expliqué dans l'annexe B.4.

Nous avons étudié cinq exploitations du causse Méjan qui ont mené à la description d'une trentaine de cas sur cette région. Ces premiers cas ont servi de base pour l'exploration de RÀPC avec notre application. Ici aussi la saisie des cas directement par l'expert agronome, plutôt que sous la dictée par l'informaticien, a montré une volonté, de la part de l'expert, d'exprimer plus de détails notamment en ce qui concernait les besoins des animaux ou encore vis-à-vis des problèmes d'embroussaillement. De plus, l'expert nous a déclaré que la saisie lui rendait évidente l'énoncé de certaines ressemblances entre des exploitations qu'il n'avait pas vues auparavant.

7.2.3 Le RàPC dans ROSA

Pour tester le RÀPC avec notre système, les agronomes nous ont proposé l'étude du nouveau chorème présenté figure 7.4. C'est une exploitation du causse de Sauveterre assez compacte. Seuls, un pacage pour les agnelles et un champ sont de l'autre côté de la route par rapport au siège de l'exploitation, tout en étant proches. Deux grands parcours incluant des blocs de champs et longés par un bois sont disposés en contiguïté.

Les agronomes ont saisi les premiers éléments du graphe de l'organisation spatiale de l'exploitation. La saisie reste assez laborieuse car on définit pas à pas les sommets et les arcs à l'aide d'une série de dialogues à renseigner. Une proposition pour saisir directement les sommets du graphe dans la vue "Gos" est à l'étude et sera implantée dans

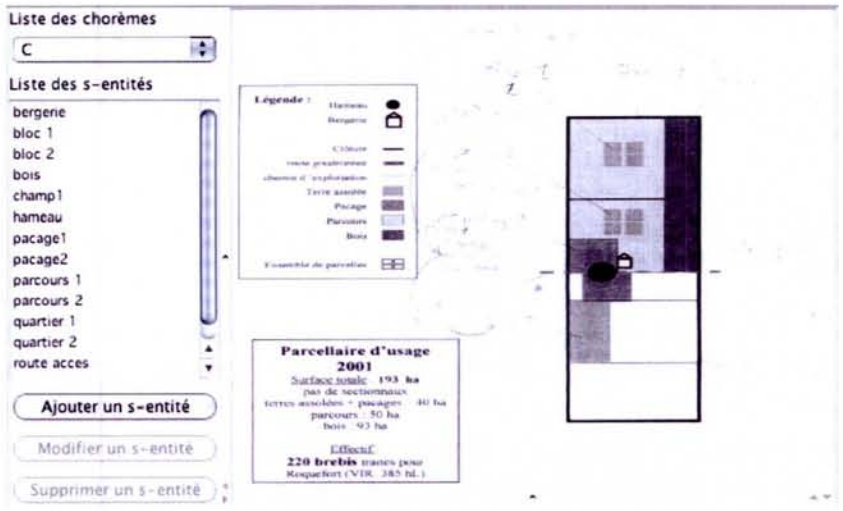


FIG. 7.4 – Exemple d'un nouveau chorème étudié.

la prochaine version de ROSA. Ce mode de saisie facilitera notamment l'établissement des arcs par sélection directe des sommets entités mis en relation par le sommet relation édité.

Le graphe sommaire présenté figure 7.5 a donc été saisi, et a servi de cible pour la remémoration. Ce graphe décrit partiellement le chorème présenté sur la figure 7.4. D'une part, un champ se trouve être difficile d'accès car une route le sépare du siège d'exploitation ce qui demande une attention particulière pour y mener le troupeau. D'autre part, deux pacages (parcours clôturé) sont placés autour du siège d'exploitation. Un seul pacage est considéré dans le graphe saisie pour des raisons de simplification. Enfin deux parcours, dont un lointain du siège sont longé par un bois et contiennent des blocs de cultures au milieu des zones en herbe. Encore une fois, un seul des deux parcours est considéré dans le graphe. C'est le parcours le plus lointain qui pose le plus de problèmes vis-à-vis de l'embroussaillement et qui est le plus intéressant du point de vue de la pratique de conduite du troupeau qui est représenté dans le graphe.

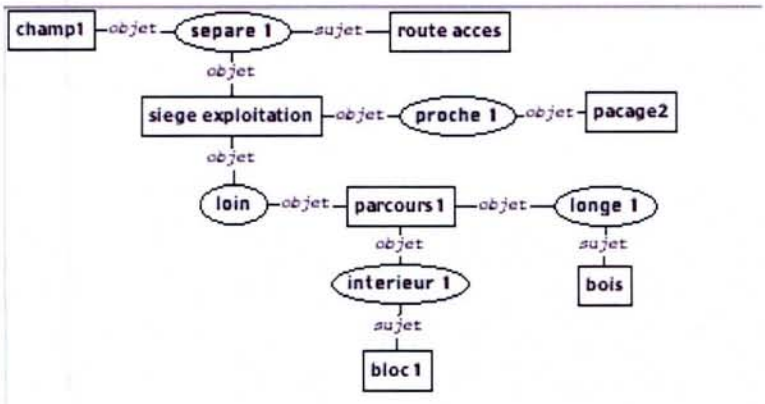


FIG. 7.5 – Graphe d'organisation d'une nouvelle exploitation que l'on cherche à analyser.

Nous recherchons donc les modalités de fonctionnement les plus vraisemblables liées à ces trois parties de graphe à l'aide de notre système ROSA, d'une part à l'aide d'une classification dure de ce graphe par rapport aux cas de la base de cas, puis d'autre part en mettant en œuvre la classification "semi dure".

7.2.3.1 Classification dure

Dans un premier temps, chaque graphe source de la base de cas est doté d'un graphe index qui est son graphe-concept. Sur la figure 7.6 le graphe SIÈGE–PROCHE–PARC-CLOS représente l'index d'un cas associé à une des exploitations du Méjan précédemment étudiées. Aucune opération de généralisation n'est alors effectuée et seule l'instanciation du cas est réalisée. Le graphe-concept obtenu est alors le graphe le plus spécifique, sur les hiérarchies d'occupation du sol et de relation, pour le graphe source. Le graphe cible réduit à `siège–proche–pacage2` sur la figure 7.6 est classé sous cet index.

Comme on peut s'y attendre, très peu de résultats de remémoration sont obtenus. L'exemple de la figure 7.6 est la seule remémoration obtenue sur notre base de cas. Ce type de résultat est prévisible dans la mesure où chaque sommet des graphes de la base de cas est typé par un concept feuille de la hiérarchie des concepts du domaine et donc extrêmement spécifique. Chaque cas source représente un très faible espace du domaine des possibles. La probabilité de trouver un graphe identique est aussi très faible. Seules les situations très habituelles sont remémorées.

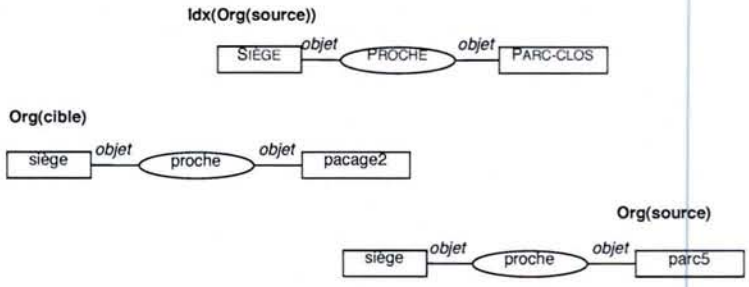


FIG. 7.6 – Classification dure du Gos de la figure 7.5 sur la base de cas. Une partie du `gos(cible)` est appariée à un cas de la base de cas.

L'explication associée au cas remémoré est : « L'agriculteur utilise ce grand parc (ressource) et proche du siège (surveillance) pour y faire pâturer ces béliers ». Ce cas correspond à une situation fort habituelle dans pratiquement toutes les exploitations sur le causse. Les éleveurs utilisent la proximité du siège pour assurer la surveillance de leurs animaux à forts besoins. L'explication peut être presque recopiée telle quelle. Il suffit de généraliser le terme de bélier par un concept décrivant des "*animaux à forts besoins*".

Le cas remémoré selon cette classification dure est validé par les agronomes, mais l'on voudrait en trouver d'autres pour expliquer les autres parties du graphe cible.

7.2.3.2 Classification "semi-dure"

Dans ce second cas de classification, nous envisageons la construction à la volée de graphes index plus généraux permettant de remémorer plus de cas.

En appliquant la remémoration semi-dure, présentée au chapitre 5.4.3, sur l'exploitation cible, nous avons obtenu six appariements entre les cas sources et le Gos de la figure 7.5 avec un seuil de similarité entre deux concepts supérieur ou égale à 0,6. Ce seuil vaut 1 pour des concepts identiques et 0 lorsque le concept commun est la racine de la hiérarchie. La limite de 0,6 a été fixée de façon expérimentale ; les concepts sont assez proches dans la hiérarchie pour être comparés.

Nous avons retrouvé le même appariement que dans le cas de la classification dure avec une similitude parfaite de 1, mais le cas source retrouvé est aussi associé à une autre partie de la cible. Parmi les cinq autres cas sources remémorés, un cas se retrouve être apparié deux fois dans cible (cible 3 et cible 4 sont appariées au même cas source) et trois cas sources n'apparaissent qu'une seule fois avec des similitudes globales comprises entre s=0,79 et s=0,93. La figure 7.7 présente les cinq nouveaux appariements obtenus associé avec leur cas source associé.

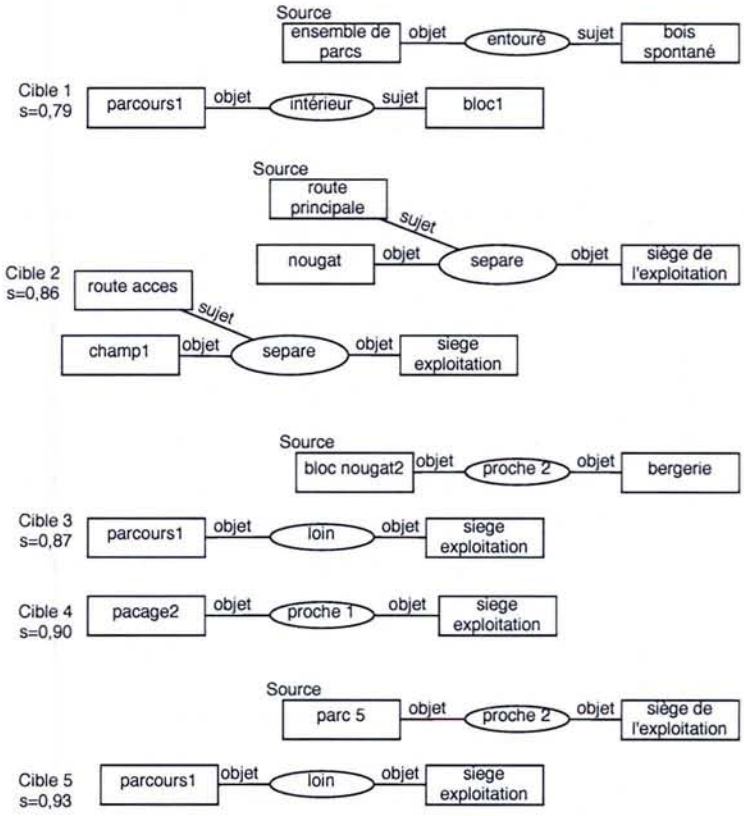


FIG. 7.7 – Les cinq nouveaux cas appariés à l'aide de la classification "semi-dure" avec leur indice de similarité.

7.2.3.3 Des cas correctement remémorés

Les cas cible 2 et cible 4 sont considérés comme acceptables. Dans le premier cas (cible 2), l'explication du source : *"Le nougat séparé par la route est utilisé quand l'exploitant a le temps de déplacer le troupeau de l'autre côté de la route"* peut être recopiée en partie. En effet les deux cas cible et source expriment une difficulté pour le déplacement du troupeau et nécessitent dans les deux cas la présence de l'agriculteur pour la surveillance du troupeau lors de l'utilisation des parcelles.

Pour le deuxième cas (cible 4), la similitude est très bonne. L'explication source est : *"Les lots d'animaux à forts besoins au printemps utilisent les surfaces en herbes du bloc nougat qui sont fertilisés (ressources), clôturés (sans garde) et proche de la bergerie (déplacements limités) alors que les champs sont clos pour être cultivés et récoltés. Après la récolte, en été (céreale) et en automne (foins), les champs sont ouverts et les animaux à forts besoins utilisent toute la surface du bloc nougat clôturé (sans garde). les lots d'animaux à forts besoins (au printemps) se succèdent dans les parties de blocs"*. Cette explication est acceptée pour la notion d'animaux à forts besoins en remplaçant le "bloc nougat" de la source par le `pacage2` dans la cible. Cependant les informations complémentaires dans la source, bien que très probables, doivent être vérifiées pour être transposées.

7.2.3.4 Cas très proches et pourtant très différents

Pour le cas cible 1, la différence sémantique entre les relations est grande alors que les concepts sont proches dans la hiérarchie. L'explication qui est : *"La présence de bois spontanés autour des parcs induit un risque de diffusion de ligneux dans les parcs"* ne convient pas. L'explication attendue ici correspond à une attraction exercée par le `bloc1` pour que les brebis pâturent le `parcours1`. De tels cas existent dans la base de cas, mais dans des structures spatiales plus complexes que celles évoquées ici, raison pour laquelle elles n'ont pas été remémorées.

De même, pour les cas cible 3 et cible 5, la remémoration n'est pas valable. Dans la hiérarchie des relations que nous avons élaborée, les concepts LOIN et PROCHE sont des sous-types de la relation binaire DÉCONNECTÉE. Ils sont relativement proches dans la hiérarchie, et fortement différents dans leur sémantique.

Pour éviter de tels écueils, trois solutions peuvent être envisagées :

- Imposer un seuil de similitude entre les types plus élevés,
- Augmenter la profondeur de la hiérarchie rendant les concepts suffisamment différents par rapport au seuil de similitude acceptable,
- Ne plus s'appuyer sur la hiérarchie ou en limiter fortement l'influence dans le calcul de similarité.

Les deux premières solutions ne sont guère satisfaisantes. Le seuil peut être ajusté mais correspond toujours à une limite un peu arbitraire. Deux seuils pourront être envisagés pour intervenir différemment entre concept de région et concept de relation. Décrire plus richement la hiérarchie peut aussi s'avérer intéressant pour préciser plus avant les concepts et leurs propriétés, mais là encore ce genre de manipulation peut rendre la hiérarchie plus complexe que nécessaire et introduire dans la définition de connaissances du domaine

une part d'arbitraire, sans compter qu'il faudra que l'expert soit vigilant vis-à-vis de la comparaison de toutes les paires de concepts, ce qui est exclu pour des hiérarchies, même de taille modeste.

7.2.3.5 Cas non trouvés que l'on aurait souhaités

Ce problème est le complémentaire du précédent. Des appariements ne sont pas proposés alors que pour certains cas on aurait voulu qu'ils ressortent. Par exemple, lorsqu'un champ se trouve au milieu d'un parcours, il joue un rôle d'attracteur pour les brebis et permet des pratiques pastorales de répartition de la pression de patûrage sur les surfaces alentours. Selon ce point de vue d'attrait, un point d'eau situé au milieu d'un parcours a le même rôle. Pourtant, un point d'eau et un champ sont deux concepts non comparables dans notre hiérarchie qui représente le point de vue d'occupation du sol.

Aussi, nous proposons de garder la hiérarchie d'occupation du sol d'une façon générale et pour des points de vue particuliers, attachés au cas décrit, s'appuyer sur des listes de concepts considérés comme étant proches en "court circuitant" en quelque sorte la hiérarchie de départ.

7.3 Usages du système

Par le travail en inter-disciplinarité, chorèmes et graphes ont évolué. Les chorèmes ont gagné en expressivité par l'adoption de règles de représentation. De même les graphes se sont stabilisés et les hiérarchies de connaissances ont évolué au gré des besoins de représentation. L'expérience d'acquisition de connaissances sur les exploitations du Sauveterre a permis de stabiliser les symbolisations graphiques. Des agronomes d'horizons différents ont, à cette occasion, trouvé des symboliques graphiques qu'ils ont formalisés et qui leur permet de parler un même langage autour des chorèmes.

L'utilisation de ROSA pour analyser une exploitation du Sauveterre en comparaison avec des exploitations du Méjan nous montre que nombre de connaissances acquises sur une région sont transposables dans une autre (généricité). Cependant, il faut considérer des exploitations de mêmes types. En effet, nous n'avons pas considéré comme comparables les exploitations de la Lorraine et celle des Causses. D'ailleurs, nous utilisons des hiérarchies de connaissances différentes pour les deux régions. L'enregistrement des données et connaissances étant réalisé dans des fichiers différents. L'application ROSA ne traite qu'une base de connaissances à la fois, elle est sélectionnée à l'ouverture de l'application par sélection du fichier à lire.

Graphe et chorème ne représentent pas la même chose. Le graphe est un moyen d'expliciter le contenu du chorème, il est manipulable par ROSA. Le chorème rend mieux compte du territoire surtout grâce à son caractère graphique - qui permet une visualisation et une compréhension intuitive de l'ensemble des objets constituant le territoire d'une exploitation agricole mais aussi de son environnement. Le graphe est tout de même une simplification du chorème. Il peut cependant avoir une richesse et une complexité assez grande ; il suffit de représenter les détails qui peuvent avoir une importance.

Alors, se pose le problème de la gestion des connaissances entrées dans le graphe.

Le niveau de détail peut être important. La densité des informations manipulées avec le graphe peut rendre le temps de calcul pour les appariements démesuré. Très rapidement, on peut atteindre un phénomène de sur-apprentissage des connaissances sur l'exploitation décrite et rendre peu visibles les grandes tendances de sa gestion au profit de remarques de détail. La limite dans le niveau de détail représenté est fonction du sens que l'information a pour les interlocuteurs.

Nous nous sommes focalisés sur les connaissances d'ordre spatial. Ce qui fait la richesse de cette démarche d'analyse du fonctionnement des exploitations tient dans l'approfondissement de l'étude des relations spatio-fonctionnelles. C'est entre autre pourquoi le chorème a toute son importance. Mais de nombreuses autres connaissances pourraient être apportées en prenant en compte d'autres aspects de l'exploitation comme la description des usages, la quantification des surfaces et leur relativité. La temporalité et les épisodes climatiques devraient aussi être mieux prises en compte tant la gestion du troupeau dépend des productions et de leur cycle mais, peut aussi dépendre de phénomènes climatiques particuliers pour sa protection et son alimentation. Aussi, la prise en compte de ces autres connaissances nécessitera la gestion de nouvelles hiérarchies de relations fonctionnelles.

Intégrer ces connaissances n'est pas une chose évidente. Nous pensons à ce niveau qu'un travail plus en profondeur notamment avec les linguistes et les psychologues devrait permettre de mieux maîtriser les capacités de représentation des connaissances exprimées au cours des enquêtes sous forme de discours ou de questions réponses. Dans ces conditions, il devra être possible de décrire d'autres types de graphes tels que des diagrammes temporels ou/et de flux de consommation/production ou d'autres formes à inventer.

Nous avons travaillé sur des exploitations très (Méjan) ou assez (Sauveterre) connues. Pour toutes ces exploitations étudiées, les agronomes possédaient des connaissances fonctionnelles de bases qui leur permettaient d'élaborer des chorèmes d'exploitation avec un minimum d'informations. Or, pour l'approche d'une nouvelle exploitation, seules sont accessibles dans un premier temps des connaissances d'ordre purement spatial (analyse de photos aériennes, cartes, parcellaires). Nous avons manipulé des connaissances d'ordre spatio-fonctionnel pour décrire les organisations spatiales. Est-il alors possible d'inférer des explications fonctionnelles sans ces connaissances ? Vraisemblablement oui dans la mesure on l'on reste dans le domaine de validité du modèle, car certaines relations fonctionnelles peuvent être inférées à partir de relations spatiales.

Une autre possibilité serait d'utiliser ce système « à l'envers » i.e. appliquer un raisonnement inverse en prenant en entrée des modes de fonctionnement et décrire en sortie des types de paysages obtenus par ces pratiques. On mettrait vraisemblablement en évidence que le même type de fonctionnement peut être atteint par des organisations spatiales différentes, ce qui correspond aux marges de manœuvre des agriculteurs pour répondre à un enjeu donné, comme celui de l'embroussaillement.

Chapitre 8

Conclusion et perspectives

Au terme de ce travail, nous pouvons dresser un bilan selon deux plans avant d'en tracer les perspectives. Tout d'abord, nous avons réalisé un système de RÀPC concret avec la définition des connaissances nécessaires à l'élaboration de raisonnements. Ensuite, ces réalisations, en terme d'application et d'organisation des connaissances, nous ont permis de dégager des enseignements sur le développement d'un système de RÀPC.

8.1 Réalisations

Nous avons présenté le développement d'un système à base de connaissances fondé sur l'utilisation du raisonnement à partir de cas. Le développement d'un tel système nécessite la manipulation de connaissances d'experts ainsi que la modélisation de raisonnements. Nous avons donc, dans un premier temps, étudié les différents formalismes utilisés en intelligence artificielle pour représenter des connaissances dans un système à base de connaissances. Puis, nous avons envisagé le raisonnement à partir de cas en milieu hiérarchique. Après avoir présenté le problème posé par les agronomes, au chapitre 3, nous avons étudié le développement de ROSA en passant par des phases d'acquisition et de modélisation aussi bien pour la représentation des connaissances que pour la modélisation des raisonnements.

8.1.1 Un modèle des connaissances agronomiques

Les connaissances agronomiques manipulées au cours de ce travail sont de deux ordres. D'une part, nous avons travaillé sur la représentation graphique utilisée par les agronomes pour synthétiser leurs connaissances sur les exploitations agricoles étudiées. D'autre part, nous leur avons fait exprimer leurs connaissances générales sur leur domaine d'application.

Au chapitre 4, nous avons montré comment nous avons traduit les chorèmes, synthétisant les connaissances sur les exploitations agricoles, sous forme de graphes de description des organisations spatiales. Nous avons aussi fait l'acquisition des connaissances du domaine, sous la forme de concepts que nous avons organisés dans des structures hiérarchiques, tant pour les entités qui représentent des occupations du sol que pour les relations spatiales entre elles. Nous avons aussi décrit d'autres connaissances utiles dans

le cadre du raisonnement à partir de cas.

8.1.2 Un modèle de RàPC sur les structures spatiales

Le travail de cette thèse permet de répondre en partie à deux des questions à l'origine de notre démarche. D'une part, nous avons montré la possibilité de manipuler des structures complexes représentées par des graphes d'entités et de relations. D'autre part, nous pouvons appliquer le modèle de RÀPC proposé dans [72]. Au chapitre 5, nous présentons comment nous pouvons manipuler les graphes dans un milieu hiérarchique à l'aide des logiques de descriptions. Puis, nous définissons des transformations applicables sur les graphes : inférences sur les relations spatiales, généralisation, pour étendre les capacité de remémoration en utilisant les chemins de similarité. Nous nous écartons du travail sur les molécules chimiques avec RESYN/RÀPC [72] dans le sens où les graphes de molécules sont plus simples : trois types de liaison, pas de sommet de type relation. De plus, notre domaine d'application est moins structuré que le domaine de la chimie et en ce qui concerne les solutions aux problèmes du RÀPC, nous travaillons sur des explications qui représentent d'autre types d'objets que ceux manipulés au niveau des graphes.

8.1.3 Un système utilisable

Ces modèles de connaissances et de raisonnement ont mené au développement de notre système ROSA que nous avons pu utiliser dans le cadre de l'étude de nouvelles exploitations comme nous l'avons présenté au chapitre 7. Nous avons pu saisir la description de l'organisation spatiale de plusieurs exploitations agricoles sur les deux régions étudiées par les agronomes avec qui nous avons travaillé. Les premiers cas de la base de cas ont été exprimés et enregistrés pour expliquer les fonctionnements observés en fonction des organisations spatiales considérées.

Les premiers tests ont révélé les possibilités d'utilisation de notre système, mais ont aussi révélé de nombreuses améliorations et évolutions possibles. Notamment, en l'état actuel, le contrôle des raisonnements de remémoration est à la charge de l'utilisateur tout comme l'enregistrement des nouveaux cas à entrer dans la base de cas.

8.2 Enseignements

Du développement de notre système, nous pouvons tirer des enseignements intéressants notamment concernant à la phase d'acquisition de connaissances, primordiales dans un système à base de connaissances, mais aussi sur la manipulation de structures complexes.

8.2.1 Acquisition de connaissances

Il arrive trop souvent que l'étude de systèmes à bases de connaissances sur des problèmes académiques ne puisse ensuite être adaptée à des applications concrètes. Nous avons pu expérimenter la difficulté de faire émerger les connaissances implicites utilisées dans leurs

raisonnements par les experts. Le travail inter-disciplinaire et la mobilisation d'outils graphiques – chorèmes des agronomes et graphes des informaticiens – ont permis d'élaborer des bases de connaissances à même de pouvoir être utilisées sur des cas concrets même si l'acquisition de connaissances est un processus jamais totalement terminé.

Nous avons montré l'intérêt de l'utilisation de formalismes graphiques pour permettre la communication entre des personnes venant d'horizons différents. Les graphes sont des bons "outils graphiques" pour représenter des structures spatiales qui sont "lisibles" par les agronomes. Nous avons aussi constaté que la confrontation des outils apporte plus que la somme de leurs intérêts respectifs. C'est ce que les psychologues et les linguistes, qui nous ont observés durant les phases d'acquisition, nous ont montré. Ils ont révélé une co-construction, par l'ensemble des acteurs, de nouveaux concepts modélisés pour décrire et préciser des connaissances plus ou moins implicites dans l'esprit des experts du domaine. Le travail d'acquisition de connaissances pour un système informatique se doit donc d'aller au cœur des connaissances, pour révéler la richesse des connaissances implicites que chaque expert utilise pour mener ses raisonnements. Ce travail a permis de faire avancer la réflexion des agronomes tant sur leur modèle graphique (les chorèmes) que sur la façon, les conditions de comparaisons des différentes situations. Au terme de cette recherche, les représentations chorématiques ont fortement évolué, les graphes ont été précisés.

8.2.2 Les logiques de descriptions pour la manipulation des graphes

Pour notre travail, nous avons exploré les formalismes de représentation de connaissances et notamment le formalisme des logiques de descriptions. Ce dernier permet la représentation de connaissances avec une forte expressivité. D'autre part ce formalisme permet, grâce à la subsomption, des raisonnements classificatoires intéressants. Ces mécanismes de classification sont utilisés comme base de l'opération de remémoration dans le cadre du raisonnement à partir de cas. En revanche, nous nous sommes heurtés au problème de la manipulation des graphes dans le cadre des logiques de descriptions. En effet, le système RACER n'offre pas la construction de contrainte sur les concepts avec un opérateur du type "same-filler" comme le propose P. Lambrix pour la représentation d'objets composites [51].

8.2.3 Tests de prototypes

D'un point de vue pratique, nous avons développé un outil de manipulation de graphes dans un environnement de connaissances hiérarchisées. Notre application œuvre sur des structures complexes : les graphes d'organisation spatiale, en utilisant le formalisme des logiques de descriptions et la classification. Le raisonnement à partir de cas permet de développer rapidement une première base pour montrer la faisabilité des raisonnements envisagés et offre des capacités d'apprentissage supervisé régulier et continu. Les prototypes de RÀPC peuvent voir le jour rapidement, ce qui permet de guider leur développement en relation avec l'avancée des formalisations des connaissances et raisonnement d'expert. De plus, l'utilisation de prototypes permet de déterminer des clés pour une stratégie de remémoration, menant vers une structuration de la base de cas.

8.3 Perpectives

8.3.1 Représentation et adaptation des explications

Nous ne sommes pas arrivés au terme du développement d'un système complet de raisonnement à partir de cas, notamment il y manque la possibilité d'adaptation des explications. Une première approche de la description des explications sous forme de graphe a été menée. Les sommets sont alors des parties d'explication qui peuvent être rattachés aux sommets du gos auquel ils font référence. Il faudra notamment développer des hiérarchies pour le domaine des explications telles qu'une hiérarchie des animaux et de leurs besoins physiologiques, ou encore une hiérarchie des usages et fonctionnement des exploitations.

Un travail plus riche dans le domaine du langage et notamment du traitement automatique du langage reste à faire. Une des difficultés sera entre autre de prendre en compte la négation dans les explications qui sont à mettre en regard des opérations de transformation des gos avec suppression de sommet. C'est par exemple, le cas de l'opposition entre un agriculteur qui laisse une bande d'herbe entre un champ et une rivière face à celui qui laboure jusqu'à la rive de la rivière.

8.3.2 Stratégie

Le raisonnement mis en place actuellement est un raisonnement *ad-hoc*. Un besoin de méta-connaissances, sur les connaissances des agronomes est nécessaire pour développer des stratégies plus fines que celles mises en place actuellement. Les notions de point de vue et d'échelle reste à développer et à formaliser afin d'enrichir les capacités de raisonnement de ce système. Ces méta-connaissances devraient aussi apporter des éléments pour les raisonnements en terme d'appariement de structures différentes. Les règles de transformation des graphes doivent être développées. Mais là encore, un riche travail d'acquisition de connaissances devra être mené. De plus avec l'enrichissement de la base de cas, il faudra envisager les stratégies nécessaires à son exploration.

8.3.3 Rosa et les agronomes

Des utilisations complémentaires ont été proposées qui restent à intégrer dans un outil aux capacités plus larges. Mais dans un premier temps, la formalisation des points de vue et des connaissances d'adaptation est nécessaire pour rendre le système réellement exploitable. Ensuite, d'autres modes de raisonnements pourront être envisagés dans des versions futures pour prendre en compte des problématiques différentes comme l'écologie du paysage par exemple.

Le système proposé pourra être étendu pour enregistrer d'autres connaissances ou informations sur les exploitations, telles que les parcellaires, les données d'enquêtes, les données chiffrées. On peut alors envisager d'utiliser le RÀPC comme un nouvel index de bases de données permettant de retrouver des exploitations à partir de requêtes plus complexes que celles traditionnellement utilisées avec le langage des SGBD. On peut aussi

envisager son utilisation "à l'envers" i.e. proposer des descriptions spatiales plausibles connaissant des explications de fonctionnement données.

Plus largement, nous pensons qu'une des applications du RÀPC et plus particulièrement de la remémoration peut être envisagée pour assurer des recherches sur le web à partir de requêtes complexes. Notamment, notre système pourrait être utilisé pour construire des descriptions personnalisées sur des domaines particuliers et être couplé ensuite à un système d'analyse de pages web pour vérifier de façon plus profonde l'adéquation entre les documents retrouvés et le sujet de la recherche.

Annexes

Annexe A

Le département Sciences de l'Action et Développement

Au sein de l'INRA, ce sont les agronomes du département *Systèmes Agraires et Développement* (SAD : nouvellement converti en *Sciences de l'Action et Développement*) qui s'attachent à comprendre et à rendre compte de la diversité des fonctionnements des exploitations agricoles et de leur évolution dans ce nouveau contexte de multifonctionnalité. Depuis 1979, ce département pluridisciplinaire regroupe des agronomes dont l'objectif est de proposer des méthodologies et des contenus afin de permettre aux acteurs de terrain - individus, groupes, réseaux, institutions - de mieux maîtriser la gestion de leur environnement [26]. L'enjeu de ces travaux est de permettre la prise en compte des nouvelles attentes sociales et des nouvelles politiques territoriales par ces acteurs. Cette prise en compte implique des changements nécessitant l'apprentissage de nouvelles techniques, en interaction avec l'élaboration de connaissances nouvelles et la mise en place de nouveaux dispositifs socio-techniques qui donnent du sens aux interventions sur le milieu et les régulent. Le projet scientifique du département Systèmes Agraires et Développement est de rendre compte de ces processus, avec une référence particulière aux espaces que gèrent les systèmes techniques mis en oeuvre par les acteurs concernés.

Pendant longtemps, les agronomes se sont intéressés aux pratiques culturales au niveau de la parcelle et plus rarement au niveau de l'exploitation. Ce nouveau projet scientifique déplace le cœur des recherches de l'exploitation vers ses rapports à des filières et à des territoires. Concrètement, les agronomes du SAD s'impliquent dans ces processus de changements techniques et organisationnels, en développant des partenariats diversifiés avec des agriculteurs, mais aussi avec des collectivités territoriales et des associations de protection de la nature et de consommateurs, à différents niveaux d'organisation. Plus précisément, quatre champs thématiques ont structuré les recherches du département SAD jusqu'en 2003 :

- Production agricole et processus de qualification ;
- Formes socio-techniques de construction et de maîtrise de problèmes d'environnement ;
- Développement local et territorialité ;
- Innovation, médiations socio-techniques et développement local.

Les trois premiers thèmes privilégient l'étude des liens entre technique et social. Le qua-

trième s'intéresse à la production de connaissances liées à l'activité de recherche avec ses partenaires de terrain. Les agronomes proposent d'étudier la manière dont sont traités en situation certains problèmes d'environnement relatifs à la protection des milieux et des espèces ainsi qu'à la qualité des eaux et à la protection des sols. Pour cela, ils étudient des situations liées à la maîtrise de problèmes environnementaux selon trois entrées conjointes :

- L'étude des systèmes d'action techniques mis en place pour gérer les processus impliqués ;
- Les formalismes et les connaissances scientifiques et tacites sur lesquelles reposent ces actions ;
- Les configurations sociales et les dispositifs d'action collective au sein desquelles ces projets se conçoivent et se réalisent.

Le département SAD est divisé en treize unités réparties sur toute la France. Ce découpage offre une diversité de domaines expérimentaux représentatifs de la gamme des situations et problèmes les plus importants de différentes régions françaises tout en menant aussi des recherches au niveau européen et avec des pays en développement. En plus de ce découpage territorial, les agronomes des différentes unités peuvent faire appel aux compétences scientifiques et techniques disponibles dans les autres unités par l'intermédiaire de projets inter-unités. Ces projets, au nombre de six actuellement, ont pour but d'organiser la production de connaissances dans les différents champs thématiques du département.

A.1 L'unité de Montpellier

Les chercheurs de l'unité de Montpellier s'intéressent aux processus d'adaptation de l'agriculture. Ils étudient l'impact des innovations techniques et organisationnelles, collectives et individuelles en cherchant à répondre à quatre objectifs :

- Produire des connaissances sur la nature et la diversité des processus d'innovation en cours dans les secteurs agricole et agro-alimentaire ;
- Élaborer de nouvelles méthodes sur la conduite de recherches pluridisciplinaires menées en partenariat avec les acteurs du changement technique ;
- Construire un cadre conceptuel commun autour de trois catégories analytiques : changement technique, apprentissage et coordination ;
- Contribuer au renouvellement de l'enseignement agronomique et professionnel par l'analyse et l'accompagnement des processus d'innovation.

Pour mener ces recherches, cette unité, constituée en UMR[27], mobilise les disciplines des sciences agronomiques (agronomie, zootechnie notamment), de l'économie et de la sociologie rurales, de la modélisation spatiale.

Les Grands Causses étudiés depuis de nombreuses années est une des régions servant de terrain d'investigation pour ces recherches nouvelles. L'analyse in situ des transformations qui ont eu lieu permet une approche des processus d'innovation induits par des phénomènes sociaux (règles, politiques publiques) ou informels (réseaux professionnels, actions collectives) [59, 107]. A travers l'étude de problématiques, telles que l'embroussaillement, en relation avec les acteurs de terrain (agriculteurs, collectivités) les chercheurs

[27] Unité Mixte de Recherche.

du SAD participent aux changements en cours et élaborent, par leur travail pluridisciplinaire, de nouvelles connaissances sur la conduite de leurs recherches [23]. Ces connaissances sont diffusées en retour vers les acteurs du terrain mais aussi à travers la formation des futurs acteurs.

A.2 L'unité de Mirecourt

L'unité de Mirecourt se situe dans une région de polyculture-élevage de l'est de la France. Cette unité est aussi une station expérimentale particulière, constituée d'une exploitation laitière comportant un troupeau de 350 bovins et disposant d'une superficie de 235ha de surface agricole utile. Les travaux menés à Mirecourt visent à identifier et à modéliser les compatibilités entre activités agricoles, la qualité des produits laitiers et la préservation des ressources en eaux. Les modalités, les conditions et certaines conséquences de la dynamique des pratiques agricoles productives y sont analysées. En particulier, l'agrandissement généralisé des exploitations qui s'accompagne d'un accroissement continu des surfaces cultivées, au détriment des surfaces en herbe, pose de nouveaux problèmes étudiés à travers deux thématiques :

- Concilier élevage laitier et valorisation de l'herbe ;
- Concilier agriculture et protection des ressources en eau.

Dans le premier cas, l'objectif est de fournir des méthodes, des outils et des références aux acteurs du développement de l'élevage et de les aider à réaliser les transformations qu'ils envisagent. Les objets de recherche privilégiés sont : l'intérêt de la mixité des troupeaux laitiers, la conduite des pâturages et des surfaces prairiales à travers le projet européen GreenGrass. Ces recherches s'appuient sur l'exploitation grandeur nature de l'unité pour :

- Concevoir, qualifier et évaluer des systèmes techniques innovants ;
- Mener des expérimentations à différentes échelles de temps et d'espace ;
- Gérer et fournir des données constitutives de son patrimoine expérimental ;
- Développer des systèmes d'observation des pratiques des agriculteurs in situ ;
- Vulgariser ses acquis.

Pour la deuxième thématique, les sites de recherche sont principalement les bassins versants lorrains impliqués dans la préservation de la qualité des eaux potables et minérales à travers les opérations Ferti-Mieux, périmètre de Vittel, zone atelier du bassin de la Moselle, et le bassin de la Seine dans le cadre du projet PIREN-Seine. Les objectifs de ces travaux sont :

- La mise au point d'itinéraires techniques peu polluants ;
- La réorganisation de l'agriculture dans les bassins d'alimentation en eau potable et minérale et le suivi de l'efficacité de ces réorganisations territoriales : occupation du sol, itinéraire technique et qualité des eaux ;
- La mise au point de méthodes de modélisation spatiale de l'organisation des systèmes techniques agricoles, de l'exploitation au grand bassin versant, en passant par la petite région agricole.

Annexe B

Manuel opérateur de ROSA

L'utilisation du système ROSA se fait à travers son interface utilisateur : *ROSA-Interface*. C'est une application écrite en JAVA présentant une interface homme machine (IHM) utilisable sur la presque totalité des systèmes informatiques du marché actuel. Cette application se constitue d'une fenêtre unique dans laquelle sont installées différentes vues accessibles par des onglets comme présenté figure B.1. Une barre de menu permet d'accéder aux menus de manipulation de fichier et à la gestion des connexions avec RACER, un serveur de connaissances qui est au cœur du module de raisonnement implanté dans ROSA.

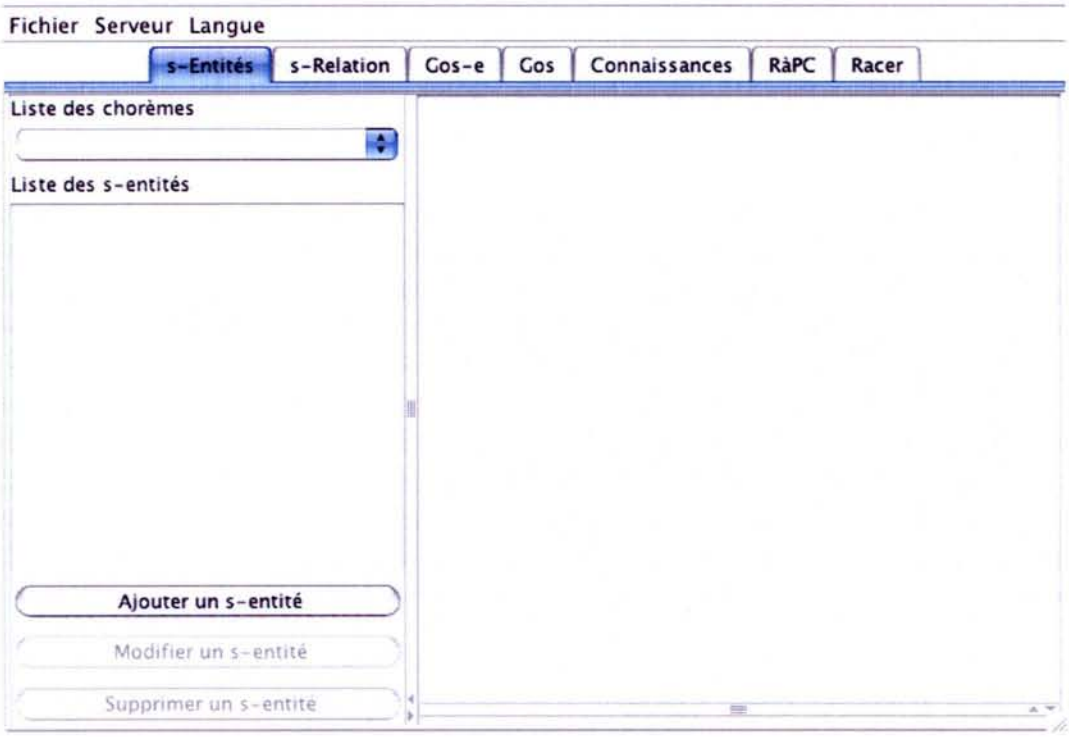


FIG. B.1 – Présentation de l'application à son ouverture.

L'interface du système ROSA est constituée de plusieurs vues décrites dans ce docu-

ment. Premièrement la vue "Connaissances" permet la saisie et la description de nouveaux concepts et rôles. Les concepts ainsi définis servent alors à typer les `s-entités` et les `s-relations` lors de la saisie des `gos`. Cette saisie de graphes s'effectue en deux étapes. Les `s-entités` sont renseignés dans la vue "s-entités", puis les `s-relations` qui permettent de construire le graphe sont renseignés dans la vue "s-relation". La vue "Cas" permet de définir les cas en sélectionnant les sommets constitutifs d'une organisation spatiale remarquable et de renseigner l'explication qui y est associée.

La vue "Gos" permet de visualiser les graphes ainsi construits. La vue "RÀPC" implante les mécanismes de raisonnement à partir de cas de ROSA. Une dernière vue "Racer" permet d'accéder directement aux fonctionnalités du système RACER qui sert de serveur de connaissances à ROSA. Pour les besoins de développement et de la maintenance, une vue "Racer" permet de connaître l'état de la base de connaissances et de dialoguer avec le serveur de logiques de descriptions du même nom.

La sauvegarde de toutes les saisies faites avec ROSA-Interface est enregistrée dans des documents au format XML (eXtensible Markup Language). La structure de ces documents est conforme à la définition du type de document spécifiée dans la DTD (Document Type Definition) qui est développée sur la base de la DTD DLML (Description Logics Modeling Language). Cette dernière permet la description des concepts et des rôles manipulés dans les logiques de descriptions. Cette abrevdtd dlml est étendue pour permettre la description des individus que sont les chorèmes, leurs `s-entités`, leurs `s-relations` et leurs `gos-e` associés.

B.1 Vue de saisie des Concepts de la base de connaissances

Cette vue (figure B.2) permet d'ajouter ou d'éditer des concepts et des rôles définis par leur nom et leur description dans le formalisme des logiques de descriptions. Lors de l'ajout d'un concept, un premier constructeur permet de préciser le type, primitif ou défini, associé au terme. Les propriétés d'un concept sont définies par sa description à l'aide d'opérateurs (and, or, all...) liant ce concept à d'autres concepts à l'aide des rôles. Dans le cas d'un concept primitif, les propriétés de ce concept sont nécessaires mais non suffisantes pour qu'une instance puisse être classée sous ce concept. En revanche, une instance réunissant les propriétés d'un concept défini sera classée sous ce concept car ses propriétés sont nécessaires et suffisantes. L'ensemble des descriptions peut alors être transmis au système RACER pour les classer les uns par rapport aux autres et vérifier la cohérence de la base de connaissances.

B.2 La saisie des Gos

La première opération pour décrire un chorème, consiste à sélectionner le chorème correspondant décrit par son nom et par un lien vers une image de type jpg ou gif qui peut alors être affichée comme support de la description des graphes qui en sont la représentation. Par une sélection dans le menu fichier, il est possible d'ajouter un

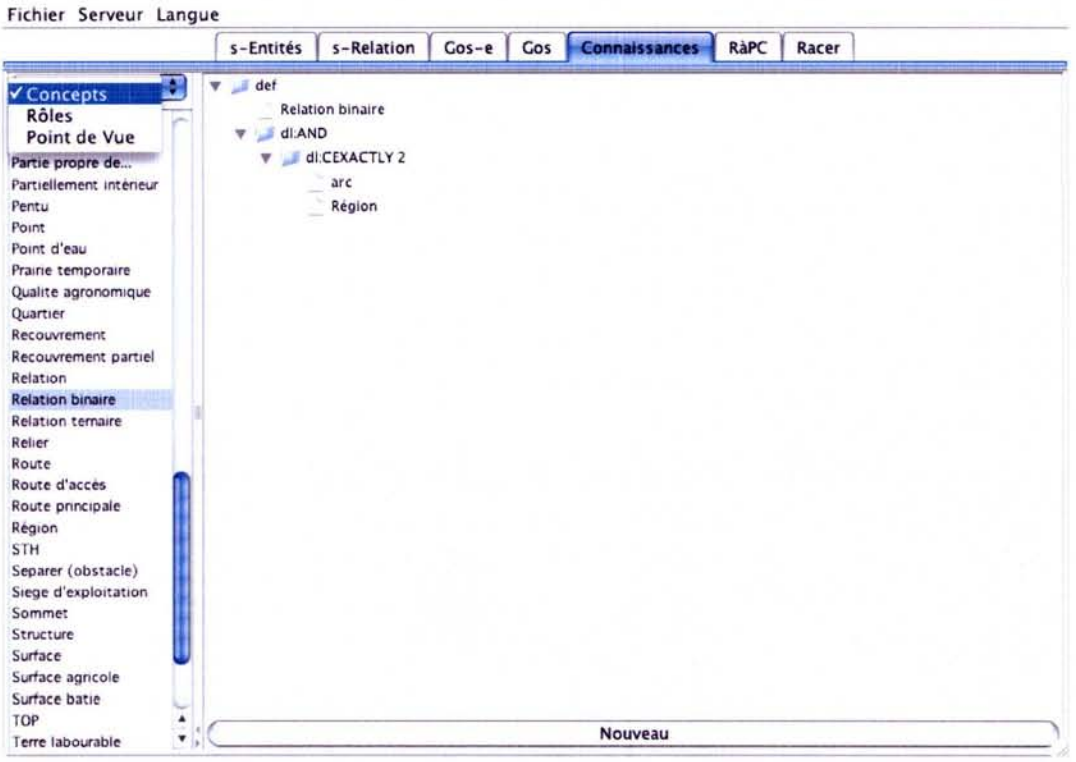


FIG. B.2 – Edition des concepts, rôles et points de vue.

chorème à la base de connaissances. Ce menu permet de définir le nom du chorème et de sélectionner l'image qui sera affichée en regard des listes d'entités et de liens.

Les vues "s-entités" et "s-relations" sont divisées en deux parties :

- à gauche, se trouve le menu déroulant qui permet de sélectionner un chorème, la liste des entités et les différents boutons ;
- à droite, se trouve l'image associée au chorème sélectionné. Si aucune image n'existe pour ce chorème, cette partie reste vide et grise.

B.2.1 Vue de saisie des s-entités

La vue "s-entités" (figure B.3) permet de sélectionner le chorème que l'on désire décrire. Le choix se fait par la sélection d'une image. De là, une liste permet d'observer l'ensemble des `s-entités` spatiaux qui y sont attachés. Un double clic sur un `s-entité` permet alors de l'éditer. Le bouton "Editer un s-entité " a le même effet. Il est aussi possible d'ajouter un `s-entité` avec le bouton "Ajouter un s-entité". Une vue s'ouvre alors de même aspect que la fenêtre de dialogue d'édition mais où le nom du `s-entité` est vide et le concept sélectionné par défaut est RÉGION. Un dernier bouton ("Supprimer un s-entité") offre la possibilité de supprimer l'entité sélectionnée.

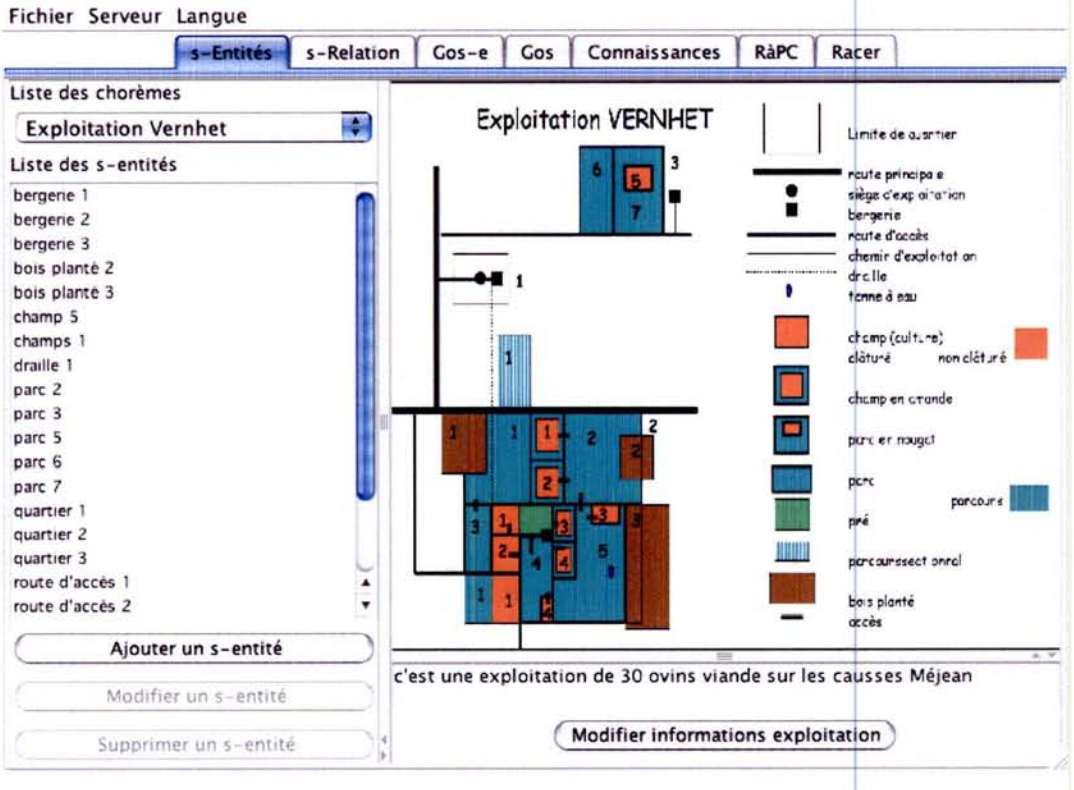


FIG. B.3 – La vue s-entité met en regard la liste des entités spatiales et le dessin du chorème d'exploitation.

B.2.2 Vue de saisie des s-relations

La vue "s-relations" est semblable à la vue "s-entités". Le choix du chorème sur lequel on veut travailler met à jour les listes dans toutes les vues et notamment la liste des entités dans la vue s-entités et la liste des relations dans la vue s-relations.

Lors d'une création, le nom est initialisé à "Nouveau s-relation" et le concept associé est "Relation". Il faut donc donner un nouveau nom et préciser un concept plus spécifique que le concept RELATION.

B.2.3 Dialogue d'édition des s-entités et des s-relations

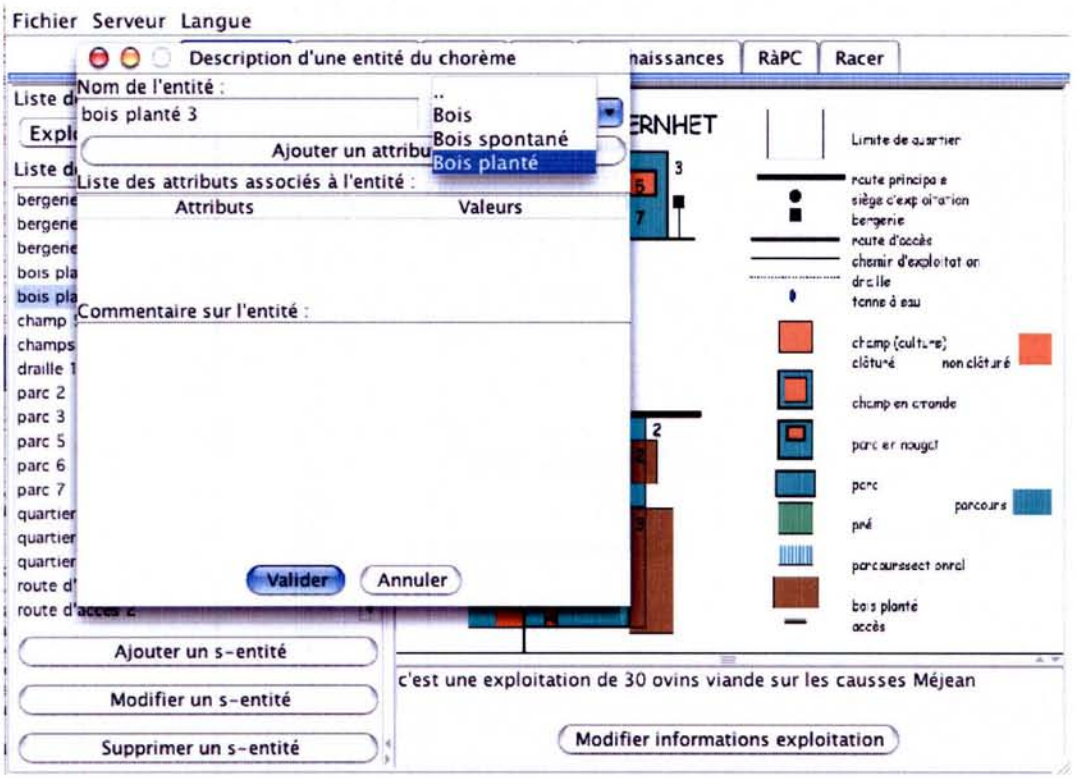


FIG. B.4 – Fenêtre de dialogue pour saisir les informations sur l'entité éditée.

Les figures B.4 et B.5 présentent les dialogues utilisés pour l'édition comme pour la création des sommets des graphes. Ces dialogues permettent d'en définir les caractéristiques, à savoir :

- Le nom.
- Le concept d'appartenance (à sélectionner ou à définir).
- Une liste de propriétés.
- Une description textuelle.

Une liste sous forme de menu déroulant permet la sélection du concept auquel rattacher l'entité. La sélection des concepts se fait suivant deux modes. Si le système est connecté au raisonneur Racer, une interaction s'établit afin de proposer, soit de généraliser le concept

soit de le spécialiser. Dans les deux cas, si la connexion au serveur de connaissances RACER est établie, le système ne propose que les concepts directement supérieur et inférieur pour limiter les sélections possibles. Sinon, en cas de déconnexion de RACER, l'ensemble des concepts du domaine est affiché (cela rend la sélection du concept moins intuitive).

Dans le cas d'un `s-relation`, il y a en plus de ces caractéristiques, une liste des `s-entités` mises en relation par le `s-relation`. Cette liste définit les arcs sortant du lien édité. Pour créer un arc, il faut appuyer sur le bouton prévu à cet effet dans le dialogue d'édition ou de création d'un `s-relation` ("Ajouter un arc"). Celui-ci ouvre une boîte de dialogue (figure B.5) qui permet de choisir l'entité à laquelle le lien est relié ainsi que le rôle que joue cette entité dans la relation (objet ou sujet par exemple). Cette boîte de dialogue est composée de deux menus déroulants. Le premier sert à choisir l'entité à laquelle l'arc est rattaché et le second permet de choisir le rôle que joue l'entité dans la relation.

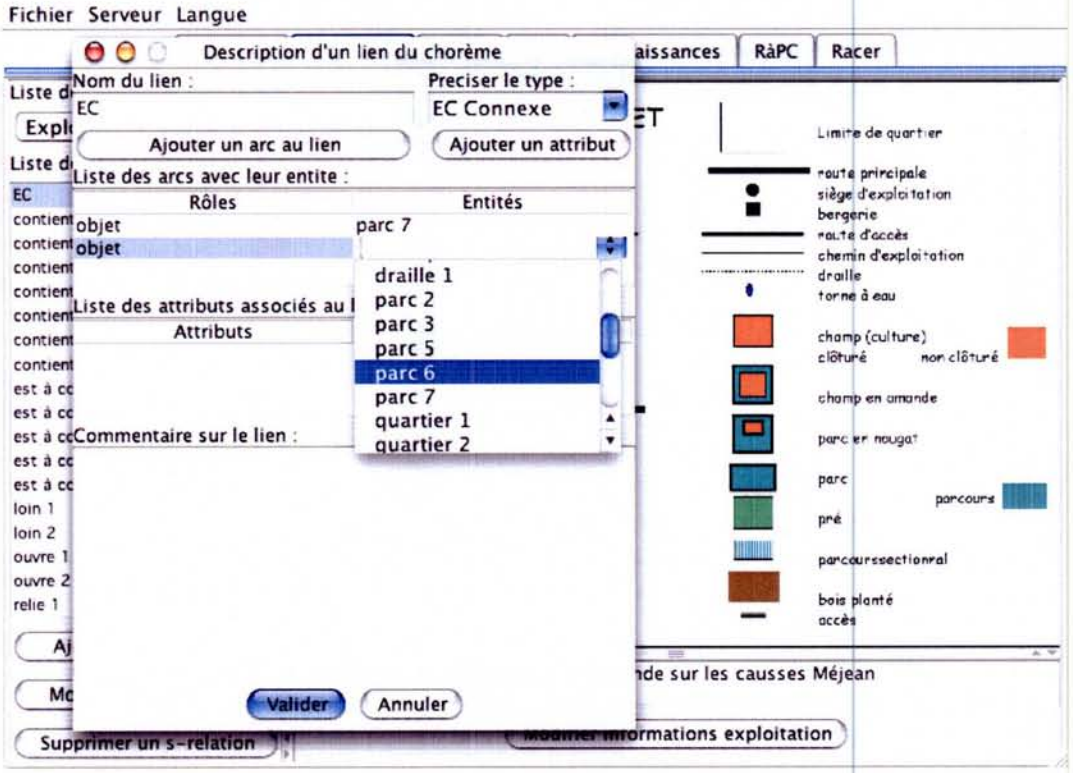


FIG. B.5 – Dialogue pour sélectionner un s-entité relié au s-relation.

B.3 La vue Gos

La vue "Gos", figure B.6, permet la visualisation des graphes des chorèmes saisis à partir des vues "s-Entités" et "s-Relations". Les sommets peuvent être déplacés dans le dessin par l'opérateur pour améliorer la lisibilité du graphe de l'exploitation sélectionnée. La forme du graphe ainsi organisée peut être enregistrée, avec le bouton "sauvegarder",

pour les visualisations ultérieures. Si l'on quitte la vue, alors que des déplacements ont été effectués, un message demande si oui ou non vous voulez enregistrer ces modifications.

Les `s-entités` sont modélisés sous forme de rectangles et les `s-relations` sont représentés par des ovales. Cette vue offre plusieurs possibilités :

- Sauvegarde des positions de chaque objet
- Visualiser un gos-e (sous-graphe) précis par sa mise en couleur (jaune) avec l'affichage de son explication associée. Un bouton permet d'activer ou de désactiver le menu déroulant qui permet la sélection des cas. L'intitulé du bouton change en fonction du choix : "Sélectionnez un cas" ou "Visualiser tout". Si aucun cas n'existe pour le chorème sélectionné, ce bouton est désactivé (figure B.6)
- Imprimer le graphe correspondant au chorème sélectionné. En appuyant sur le bouton Impression, cela lance la boîte de dialogue, du système d'exploitation, associée à l'impression

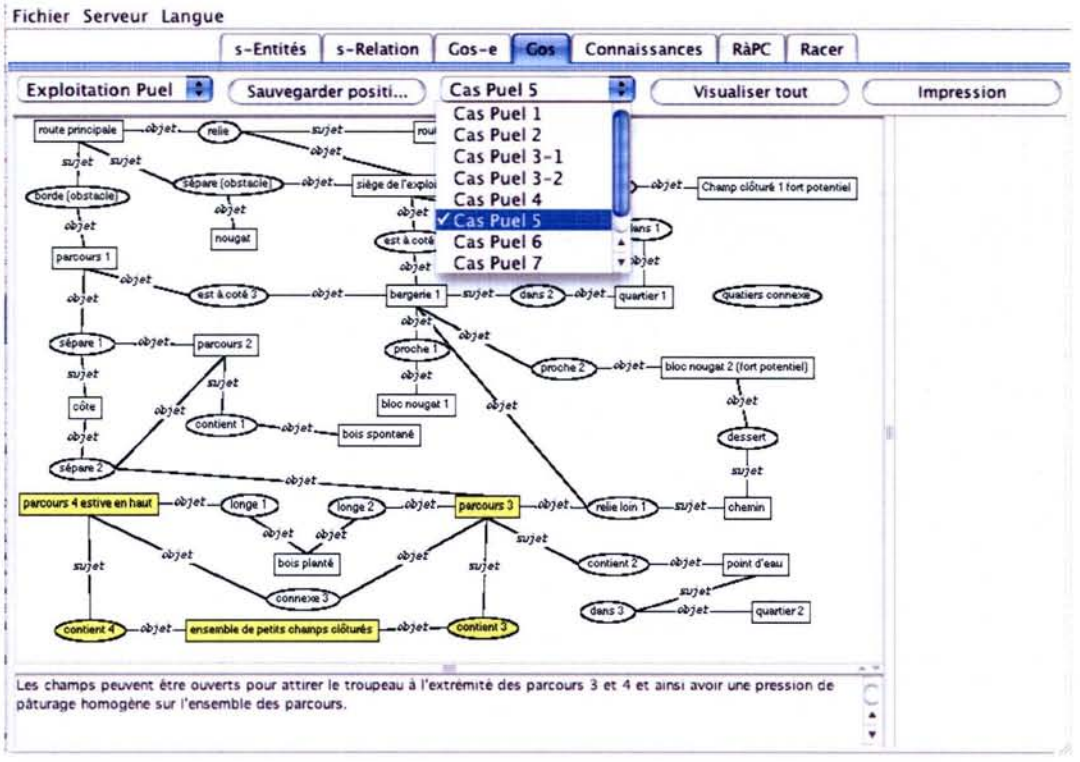


FIG. B.6 – Présentation d'un `gos` avec la sélection d'un `gos-e` mis en relief.

B.4 La vue Cas

La vue "Cas", figure B.7 permet, quant à elle, de définir les cas. Un cas se construit en sélectionnant certains sommets (`s-entités` et `s-relations`) d'un graphe représentant un chorème, pour en extraire un sous-graphe définissant une organisation spatiale remarquable à laquelle est associée une explication.

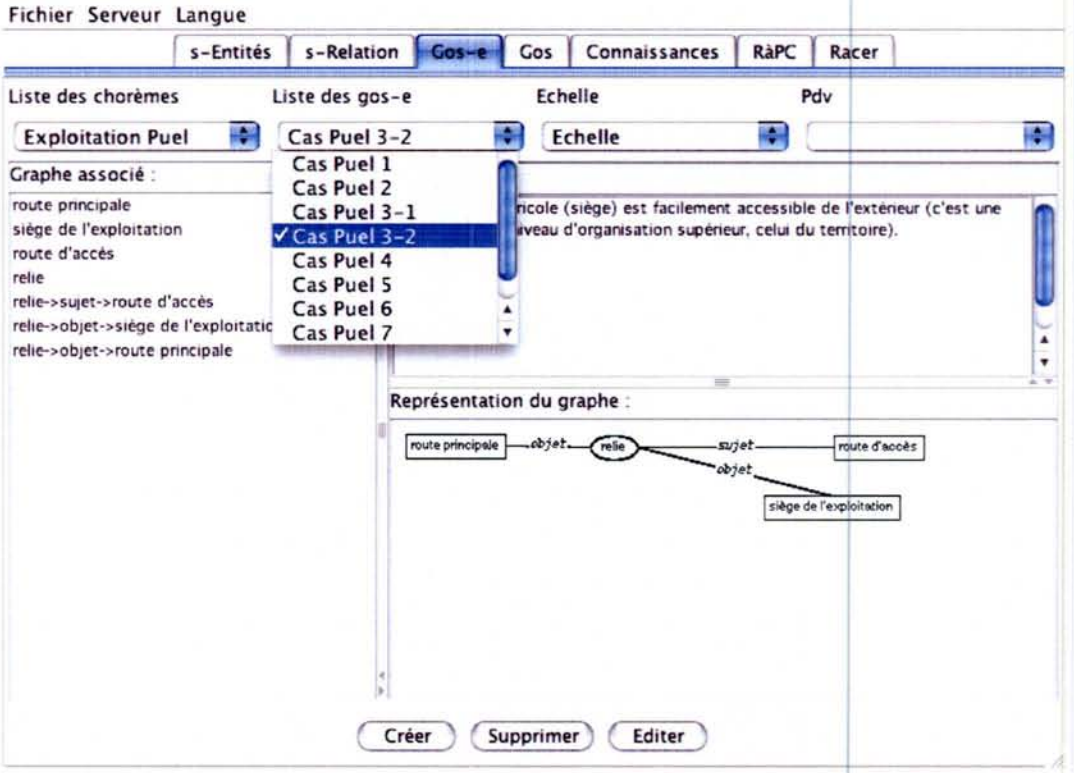


FIG. B.7 – Visualisation et édition des cas.

Cette vue offre l'accès à la liste des cas par chorème. La vue "Cas" est divisée en trois parties :

- La liste des sommets associés au cas.
- L'explication associée au cas.
- Le graphe représentant le cas.

Il est alors possible de créer un cas avec le bouton "Créer" ainsi qu'à l'aide de la boîte de dialogue correspondante (figure B.7) en sélectionnant les sommets associés au sous-graphe représentant une organisation spatiale particulière, puis en renseignant l'explication qui lui est associée.

B.5 La vue RàPC

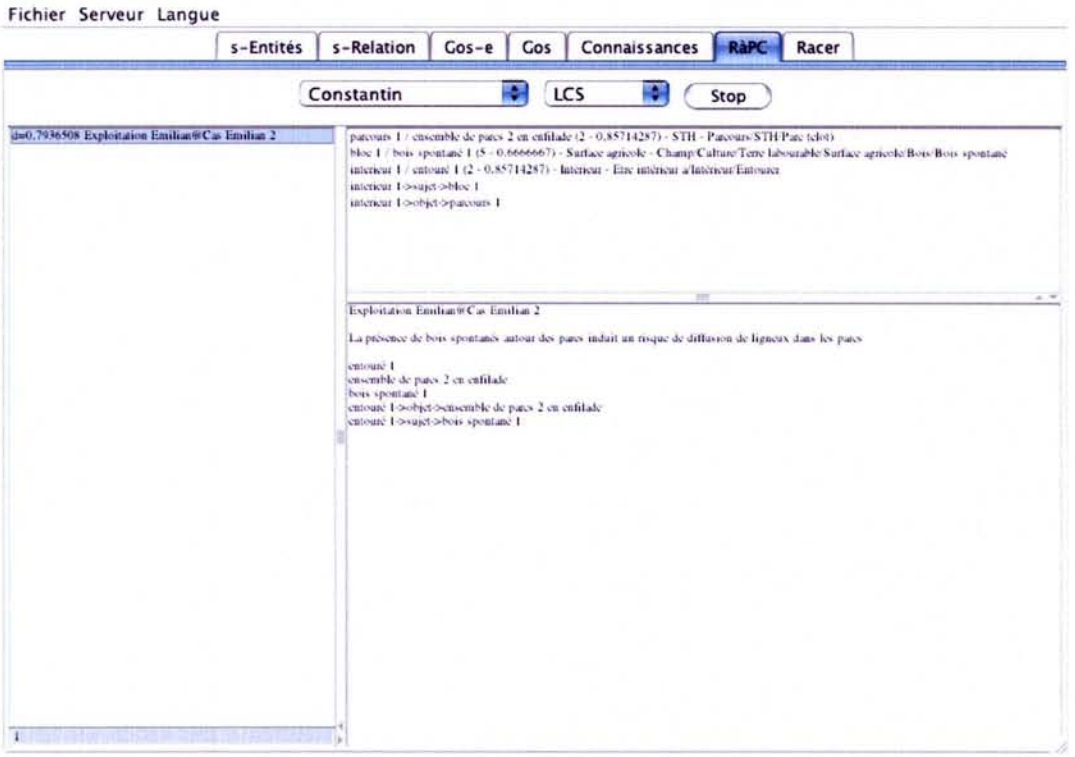


FIG. B.8 – Visualisation des appariements pour les cas remémorés.

On choisit le chorème que l'on souhaite expliquer à l'aide du raisonnement à partir de cas (figure B.8) menu "Constantin". Ensuite, on choisit le type d'appariement souhaité à partir du deuxième menu déroulant. Le choix s'effectue parmi les deux options implantées :

- Dure
- LCS

Enfin, on appuie sur le bouton "Test" pour lancer le raisonnement à partir de cas. En cours de fonctionnement, chaque graphe apparié s'affiche au fur et à mesure dans la liste située à gauche. Il sont triés par ordre de similarité décroissant. Cette vue n'autorise

actuellement que la visualisation des appariements effectués ainsi que l'explication du cas remémoré. L'enregistrement d'un nouveau cas est laissé à l'initiative de l'utilisateur via la vue "Cas".

B.6 La vue Racer

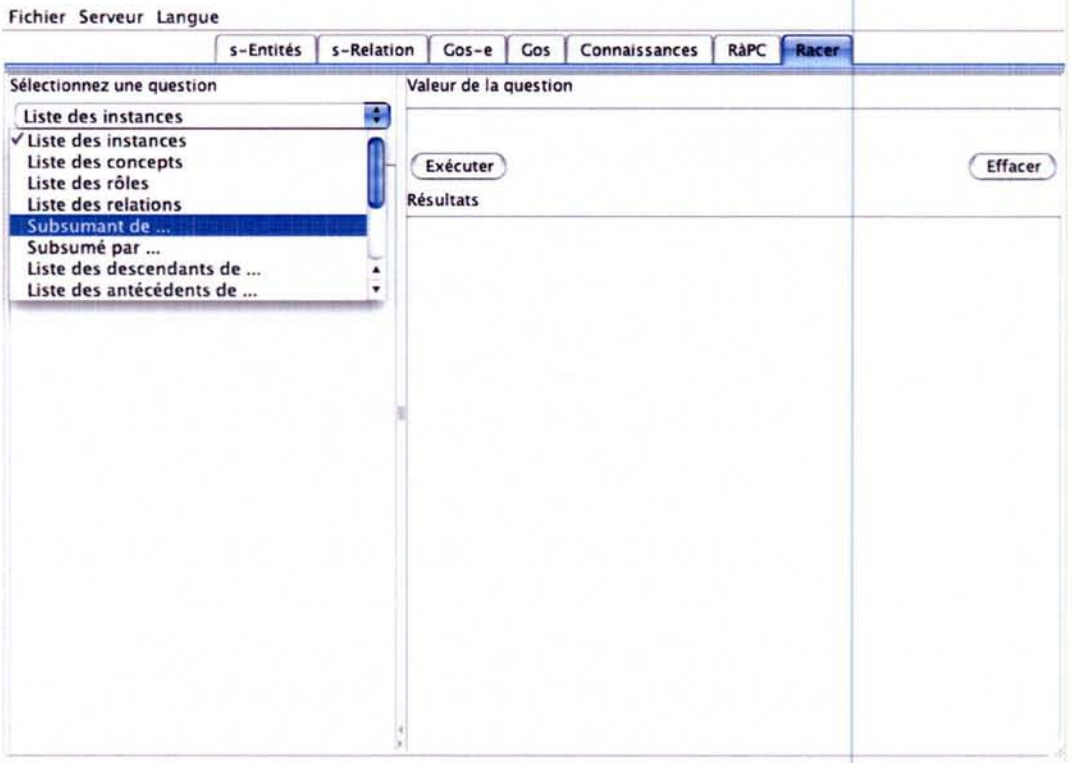


FIG. B.9 – Cette vue permet l'accès direct aux fonctions de RACER.

Dans cette vue, présentée figure B.9, il faut sélectionner une question dans le menu déroulant prévu à cet effet. Le choix s'effectue parmi les différents services proposés par le raisonneur RACER :

- Liste des instances
- Liste des concepts
- Liste des rôles
- Liste des relations
- Subsumant de ...
- Subsumé par ...
- Liste des descendants de ...
- Liste des antécédents de ...
- Instances du concept ...
- Concepts de l'instance ...
- Hiérarchie de concepts de l'instance ...

- Rôles du domaine de l'instance ...
- Rôles du co-domaine de l'instance ...

Ensuite, si la question choisie nécessite un paramètre, il est choisi dans la liste de gauche qui vient de s'afficher. Enfin, l'action sur le bouton "Exécuter" lance la requête au serveur. La liste des réponses s'affiche dans la liste située à droite. Le bouton "Effacer" permet d'effacer toutes les listes.

Bibliographie

[1] Aamodt (A.) et Plaza (E.). – Case-Based Reasoning : Foundational issues, methodological variations and system approaches. *AICOM*, vol. 1, mars 1994, pp. 39–59.

[2] Baader (F.) et Hollunder (B.). – KRIS : Knowledge Representation and Inference System. *SIGART Bulletin*, no3, 1991, pp. 8–14.

[3] Bachimont (B.). – *Le contrôle dans les systèmes à base de connaissances. Contribution à l'épistémologie de l'intelligence artificielle.* – Paris, Hermès, 1994.

[4] Baudry (J.) et Papy (F.). – The role of landscape heterogeneity in the sustainability of cropping systems. *In : Cop Science*, éd. par J. Nösberger (H. H. Geiger) et Struick (P. C.), pp. 243–259. – CAB International, 2001.

[5] Bessy-Pietri (P.), Hilal (M.) et Schmitt (B.). – Recensement de la population 1999 - Evolutions contrastées du rural. *INRA Sciences sociales*, vol. 00, n° 2, juillet 2000, p. 4.

[6] Bonin (M.) et Lardon (S.). – Recomposition des exploitations agricoles et diversification des pratiques de gestion de l'espace. *Etude Recherche Système Agraires et Développement*, vol. 33, 2002, pp. 131–148.

[7] Borgida (A.). – Description Logics in data management. *Knowledge and Data Engineering*, vol. 7, n° 5, 1995, pp. 671–682.

[8] Brachman (R. J.), McGuinness (D. L.), Patel-Schneider (P. F.) et Resnick (L. A.). – Living with CLASSIC : When and How to Use a KL-ONE-Like Language. *In : Principles of Semantic Networks : Explorations in the Representation of Knowledge*, éd. par Sowa (J. F.), pp. 401–456. – San Mateo, CA, Morgan Kaufmann, 1991.

[9] Brachman (R.J.). – "I lied about the trees" or defaults and definition in knowledge representation. *The IA Magazine*, vol. 6(3), 1985, pp. 80–93.

[10] Brassac (C). – Action située et distribuée et analyse du discours : quelques interrogations. *Cahiers de Linguistique Française*, no26, 2004, pp. 251–268.

[11] Brunet (R.). – La carte-modèle et les chorèmes. *In : MappeMonde*, pp. 2–6. – Reclus, 1986.

[12] Bunke (H.) et Messmer (B.T.). – Similarity Measures for Structured Representations. *In : Proceedings of the First European Workshop on Case-Based Reasoning (EWCBR'93), Kaiserslautern*, éd. par Richter (M.M.), Wess (S.), Althoff (K.-D.) et Maurer (F.), pp. 26–31.

[13] Calvanese (D.), De Giacomo (G.) et Lenzerini (M.). – Conjonctive Query Containment in Description Logics with n-ary Relations. *In : International Workshop on Description Logics.* – Paris, 1997.

[14] Calvanese (D.), De Giacomo (G.) et Lenzerini (M.). – Representing and Reasoning on XML Documents : A Description Logic Approach. *Journal of Logic and Computation*, vol. 9, 1999.

[15] Calvanese (D.), Lenzerini (M.) et Nardi (D.). – Unifying Class-Based Representation Formalisms. *Journal of Artificial Intelligence Research*, vol. 11, Sep 1999, pp. 199–240.

[16] Capitaine (M.). – *Organisation spatiale du territoire des exploitations agricoles. Proposition méthodologique pour interroger les conceptions des agriculteurs sur leur territoire.* – INRA Station SAD Mirecourt, Mémoire de DEA Sciences Agronomiques, ENSAIA, septembre 1999.

[17] Capitaine (M.), Lardon (S.), Le Ber (F.) et Metzger (J.-L.). – Chorèmes et graphes pour modéliser les interactions entre organisation spatiale et fonctionnement des exploitations agricoles. *In : Géomatique et espace rural. Journées CASSINI 2001, Montpellier, France*, éd. par CIRAD. pp. 145–163. – T. Libourel.

[18] Carbonell (J. G.). – Learning by analogy : Formulation and generalizing plans from past experience. *In : Machine Learning, an Artificial Intelligence Approach, Volume-1*, éd. par Michalski (R.S.), Carbonell (J.G.) et Mitchell (T.M.), pp. 137–162. – Tioga, 1983.

[19] Caron (P.). – Quels agronomes pour quels territoires ? Le terroire, un concept porteur d'intégration et de marginalisation au sein de la discipline. *Agronomes et territoires*, septembre 2002, p. 21.

[20] Chein (M.) et Mugnier (M.-L.). – Conceptual Graphs : Fundamental Notions. *Revue d'Intelligence Artificielle*, vol. 6, n° 4, 1992, pp. 365–406.

[21] Chouraqui (E.). – Le raisonnement analogique : sa problématique, ses applications. *In : Actes des journées Nationales sur l'Intelligence Artificielle.* pp. 107–117. – Aix-les-Bains, 1986.

[22] Clementini (E.), Felice (P.D.) et Hernández (D.). – *Qualitative Representation of positional information.* – Rapport technique n° FKI-208-95, Institut für Informatik, Technische Universität München, july 1995.

[23] Cohen (M. dir). – La brousse et le berger. Une approche interdisciplinaire de l'embroussaillement des parcours. *In : Espace et milieux*, p. 354. – Paris, France, CNRS Edition, 2003.

[24] Cohn (A. G.), Bennett (B.), Gooday (J.) et Gotts (N. M.). – Representing and Reasoning with Qualitative Spatial Relations About Regions. *In : Spatial and Temporal Reasoning*, pp. 97–134. – Kluwer Academic Publishers, 1997.

[25] De Giacomo (G) et Lenzerini (M.). – Description Logic with Inverse Roles, Functional Restrictions, and N-ary Relation. *In : Logics in Artificial Intelligence.* pp. 332–346. – Berlin, 1994.

[26] Deffontaines (J.-P.). – *Les sentiers d'un géoagronome.* – Editions ARGUMENTS, 1998, 359p.

[27] Deffontaines (J.-P.). – *Les entretiens du Pradel*, chap. Introduction. Une agronomie en questionnement, pp. 115–119. – Autour d'Olivier de Serres : Pratiques agricoles et pensées agronomiques, 2001, *4*.

[28] Deffontaines (J.-P.), Brossier (J.), Barbier (M.), Benoît (M.), Chia (E.), Fiorelli (J.-L.), Gafsi (M.), Gras (F.), Lemery (B.) et Roux (M.). – *Recherche Agriculture Environnement Vittel : AGREV, rapport de synthèse de fin de contrat avec la Perrier Vittel SA.* – Rapport technique, Dijon, INRA-SAD, 1996.

[29] Deffontaines (J.-P.), Cheylan (J.-P.) et Lardon (S.). – Gestion de l'espace rural, des pratiques aux modèles. *MappeMonde*, no4, 1990, p. 48.

[30] Deffontaines (J.-P.) et Lardon (S.). – Grasslands and agrarian systems. Methodological considerations on space in the management of grasslands. *Etude Recherche Système Agraires et Développement*, no16, 1989, pp. 209–218.

[31] Ducournau (R.). – Y3 : YAFOOL, the object oriented language, 1991. SEMA Group.

[32] Duvernoy (I.), Lardon (S.), Albaladejo (C.), Benoît (M.), Langlet (A.), Muhar (M.-C.) et Triboulet (P.). – Approche spatiale et fonctionnelle des relations entre activités agricoles et territoires. Construction d'une méthode de diagnostic. *In : Recherches-système en agriculture et développement rural (Sympos. Internat., Montpellier,/11/21-25).* CIRAD-SAR, pp. 230–235. – Montpellier, 1994.

[33] Egenhofer (M. J.). – Reasoning about binary topological relations. *In : Proceedings SSD'91, Advance in Spatial Databases*, éd. par Gunther (O.) et Schek (H.-J.). pp. 143–160. – Springer-Verlag.

[34] Euzenat (J.). – Preserving modularity in XML encoding of description logics. *In : Proceedings of the International Workshop on Description Logics (DL2001), Standford, USA*, pp. 20–29.

[35] Eynard (J.) et Ripoll (T.). – La détection d'analogies potentielles sans mapping est-elle possible ? *In : Acte journée RàPC 2003.* – Laboratoire de Psychologie Cognitive : UMR CNRS 6146, Juillet 2003.

[36] Fuchs (B.). – *Représentation des connaissances pour le raisonnement à partir de cas : Le système ROCADE.* – Saint-Etienne, France, Thèse de doctorat, Université Jean Monnet, 1997.

[37] Ganesan (P.), Garcia-Molina (H.) et J. (Widow). – Exploiting Hierarchical Domain Structure to Compute Similarity. *ACM Transaction on Information systemss*, vol. 21, n° 1, Jan 2003, pp. 64–93.

[38] Gautier (D.), Lardon (S.) et Osty (P. L.). – Recherche d'entités spatio-temporelles pour modéliser les dynamiques de mise en valeur de l'espace rural : des quartiers ruraux sur le Causse Méjan ? *In : Les temps de l'environnement, Journées PIREVS-CNRS, Tome 2*, pp. 125–132. – Toulouse, France, 1997.

[39] Girard (N.), Bellon (S.), Hubert (B.), Lardon (S.), Moulin (C. H.) et L. (Osty P.). – Categorising combinations of farmers' land use practices : an approach based

on examples of sheep farms in south of France. *In : Agronomie : agriculture et environnement*, pp. 435–459.

[40] Goasdoué (F.), Lattès (V.) et Rousset (M.-C.). – The Use of CARIN Language and Algorithms for Information Integration : The PICSEL System. *International Journal of Cooperative Information Systems*, vol. 9, n° 4, december 2000, pp. 383–401.

[41] Haarslev (V.) et Möller (R.). – RACER System Description. *In : Proceedings of the International Joint Conference on Automated Reasoning, IJCAR'2001, Siena, Italy*, éd. par Goré (R.), Leitsch (A.) et Nipkow (T.). pp. 701–706. – Springer.

[42] Haarslev (V.) et Möller (R.). – *RACER User's Guide and Reference Manual.* – Université de Hambourg, nov 2001.

[43] Haton (J.-P.), Bouzid (N.), Charpillet (F.), Haton (M.-C.), Lasri (B.), Lasri (H.), Marquis (P.), Mondot (T.) et Napoli (A.). – *Le raisonnement en intelligence artificielle.* – InterEditions, 1991.

[44] Hernández (D.). – Qualitative Representation of spatial knowledge. *Lecture Notes in Artificial Intelligence*, vol. 804, 1994.

[45] Horrocks (I.). – FaCT and iFaCT. *In : Proceedings of the International Workshop on Description Logics (DL'99)*, éd. par Lambrix (P.), Borgida (A.), Lenzerini (M.), Möller (R.) et Patel-Schneider (P.), pp. 133–135.

[46] Horrocks (I.). – DAML+OIL : A Reason-able Web Ontology Language. *In : Extending Database Technology*, pp. 2–13.

[47] Horrocks (I.) et Patel-Schneider (P. F.). – DL Systems Comparison (Summary Relation). *In : International Workshop on Description Logics.* – Trento, Italy, June 1998.

[48] INRIA Rhône-Alpes. – *Troeps 1.3a Reference manual*, december 1999.

[49] Koehler (J.). – Planning from Second Principles. *Artificial Intelligence*, vol. 87, 1996, pp. 145–186.

[50] Kolodner (J.). – *Case-Based Reasoning.* – San Mateo, California, Morgan Kaufmann Publishers, Inc., 1993.

[51] Lambrix (P.). – *Part-Whole Reasoning in an Object-Centered Framework.* – Springer-Verlag, 2000, *LNAI.*

[52] Lambrix (P.), Shahmehri (N.) et Jacobsen (S.). – Querying Document Bases by Content, Structure and Properties. *In : ISMIS'99*, pp. 123–132. – Springer-Verlag, 1999.

[53] Landais (E.). – Pratiques d'élevage extensif. Identifier, modéliser, évaluer. *Etude Recherche Système Agraires et Développement*, no27, 1993, p. 389.

[54] Lardon (S.), Capitaine (C.) et Benoît (M.). – Les modèles graphiques pour représenter l'organisation spatiale des activités agricoles. *In : Représentations graphiques dans les systèmes complexes naturels et artificiels.* pp. 127–150. – Journées de Rochebrune, février 2000.

[55] Lardon (S.), Le Ber (F.), Metzger (J.-L.) et Osty (P. L.). – Une démarche et un outil pour modéliser et comparer l'organisation spatiale d'exploitations agricoles. *Revue internationale de Géomatique*, 2005.

[56] Lardon (S.), Libourel (T.) et Cheylan (J.-P.). – Représentation du changement : Conceptualiser la dynamique des entités spatio-temporelles. *Revue Internationale Géomatique*, vol. 99, n° 1, 1999, pp. 45–65.

[57] Lardon (S.) et Osty (P.-L.). – Time-space dimensions of farmer practices : methodological proposals from surveys and modelling of sheep farming. *In : European Farming and Rural Systems Research and Extension into the next Millenium.* – Volos, Greece, April 2000.

[58] Lardon (S.) et Osty (P. L.). – Diversité des modes de conduite des élevages et organisation spatiale des lieux-dits : des atouts pour mieux gérer le territoire ? *In : Colloque "Gérer les paysages de montagne, pour un développement concerté et durable", CdRom CEP Florac.*

[59] Lardon (S.), Osty (P.-L.) et Triboulet (P.). – Elevage et éleveurs du Causse Méjan (Lozère) - Dynamique de la mise en valeur et contrôle de l'espace. *In : Grands Causses - Nouveaux enjeux, nouveaux regards*, pp. 219–242. – Millau, J.-L. Bonniol and A. Saussol, fédération pour la vie et la sauvegarde du pays des grands causses édition, october 1993.

[60] Lardon (S.), Osty (P. L.) et Triboulet (P.). – Elevage et éleveurs du Causse-Méjan (Lozère). Dynamique de mise en valeur et contrôle de l'espace. *In : Grands Causses. Nouveaux enjeux, nouveaux regards. Fédération pour la Vie et la Sauvegarde du Pays des Grands Causses (Millau, 1993/10/21-23)*, éd. par Bonniol J.-L. (Saussol A. (coord.)), pp. 219–242.

[61] Le Ber (F.). – *Modélisation des connaissances et raisonnements pour l'analyse des paysages agraires à partir de données satellitaires.* – Thèse de doctorat, Université Henri Poincaré - Nancy 1, décembre 1993.

[62] Le Ber (F.), Brassac (C.) et Metzger (J.-L.). – Analyse de l'interaction experts - informaticiens pour la modélisation de connaissances spatiales. *In : IC'2002, Journées Francophones d'Ingénierie des Connaissances, Rouen.* pp. 29–38. – INSA Rouen.

[63] Le Ber (F.), Brassac (C.), Préau (J.-M.) et Metzger (J.-L.). – De la confiserie sur le Causse, ou comment concilier chorèmes et graphes. *In : Agro-tribulations*, éd. par Blanc-Pamard (C.), Deffontaines (J.-P.), Lardon (S.), Raichon (C.) et Zasser-Bedoya (S.), pp. 117–130. – INRA éditions, 2004.

[64] Le Ber (F.), Mangelinck (L.) et Napoli (A.). – Représentation de relations et classification de structures spatiales. *Revue d'Intelligence Artificielle*, vol. 13, n° 2, 1999, pp. 441–467.

[65] Le Ber (F.), Metzger (J.-L.) et Napoli (A.). – Modeling and comparing maps with graphs. *In : ECAI Workshop on Spatial and Temporal Reasoning, Lyon*, pp. 31–34.

[66] Le Ber (F.) et Napoli (A.). – The design of an object-based system for representing and classifying spatial structures and relations. *Journal of Universal Computer*

Science, vol. 8, n° 8, 2002, pp. 751–773. – Special issue on Spatial and Temporal Reasoning.

[67] Le Ber (F.), Napoli (A.), Metzger (J.-L.) et Lardon (S.). – Modeling and comparing farm using graphs and case-based reasoning. *Journal of Universal Computer Science*, vol. 9, n° 9, septembre 2003, pp. 1073–1095.

[68] Leake (D. B.) (édité par). – *Case-Based Reasoning. Experiences, Lessons and Future Direction.* – MIT Press, 1996.

[69] Leclère (M.). – *Les connaissances du niveau terminologique du modèle des graphes conceptuels : construction et exploitation.* – Thèse de doctorat, Université Sciences et Techniques du Languedoc - Montpellier 2, décembre 1995.

[70] Leclère (M.). – C-CHiC : Construction coopérative de hiérarchies de catégories. *Revue d'Intelligence Artificielle*, vol. 10(1), 1996, pp. 57–100.

[71] Leeuw (S.E. van der). – Archaeomedes, un programme de recherche européen sur la désertification et la dégradation des sols. *Natures, Sciences, Sociétés*, 1998, pp. 53–58.

[72] Lieber (J.). – *Raisonnement à partir de cas et classification hiérarchique. Application à la planification de synthèse en chimie organique.* – Thèse de doctorat, Université Henri Poincaré - Nancy 1, octobre 1997.

[73] Lieber (J.). – Recopier c'est déjà adapter : six types d'adaptation par copie. *In : RaPC 2002 : Atelier Raisonnement à Partir de Cas*, pp. 11–21. – Paris, 2002.

[74] Lieber (J.) et Napoli (A.). – Using Classification in Case-Based Planning. *In : Proceedings of the 12th European Conference on Artificial Intelligence (ECAI'96)*, éd. par Wahlster (W.). pp. 132–136. – Budapest, Hungary, 1996.

[75] Lieber (J.) et Napoli (A.). – Correct and Complete Retrieval for Case-Based Problem-Solving. *In : Proceedings of the 13th European Conference on Artificial Intelligence (ECAI'98), Brighton, UK*, éd. par Prade (H.). pp. 68–72. – J. Wiley & Sons Ltd, Chichester.

[76] Lieber (J.) et Napoli (A.). – Raisonnement à partir de cas et résolution de problèmes dans une représentation par objets. *Revue d'Intelligence Artificielle*, vol. 13, 1999, pp. 9–35.

[77] Lifran (R.). – Politiques publiques et dynamique des paysages au Sud du Massif Central. *In : Rapport final MATE, programme "Politiques Publiques et Paysages"*, p. 168. – Montpellier, INRA , UMR Lameta, 2003.

[78] MacGregor (R.) et Bates (R.). – *The LOOM knowledge representation language.* – Rapport technique n° RS-87-188, Univ. of Southern California, Information Sciences Institute, 1987.

[79] Mangelinck (L.). – *Représentation et classification de structures spatiales. Application à la reconnaissance de paysages agricoles.* – Thèse de doctorat, Université Henri Poincaré - Nancy 1, octobre 1998.

[80] Mangelinck (L.) et Le Ber (F.). – Représentation de relations et classification : Reconnaissance de structures spatiales agricoles sur images satellites. *In : Actes de*

S.C.D. - U.H.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

la conférence Langages et Modèles Objets (LMO'99). pp. 197–212. – Villefranche-sur-Mer, janvier 1999.

[81] Mariño Drews (O.). – *Raisonnement classificatoire dans une représentation à objets multi-points de vue.* – Thèse de doctorat, Université Joseph Fourier - Grenoble 1, octobre 1993.

[82] Marres (P.). – *Les Grands Causses : étude de géographie physique et humaine.* – Tours, Arrault and Cie, 1953.

[83] Masini (G.), Napoli (A.), Colnet (D.), Léonard (D.) et Tombre (K.). – *Les langages à objets.* – InterEditions, 1989.

[84] Matsuyama (T.) et Hwang (V.S.). – *SIGMA - A Knowledge-Based Aerial Image Understanding System*, chap. System Architecture and Reasoning Scheme in SIGMA, pp. 35–97. – Kluwer Academic, Boston, USA, 1990.

[85] McGuinness (D. L.) et Borgida (A.). – Explaining Subsumption in Description Logics. *In : 14th International Joint Conference on Artificial Intelligence*, éd. par Mellish (C.). pp. 816–821. – San Francisco, California, 1995.

[86] Melis (E.), Lieber (J.) et Napoli (A.). – Reformulation in Case-Based Reasoning. *In : Fourth European Workshop on Case-Based Reasoning, EWCBR-98*, éd. par Smyth (B.) et Cunningham (P.). pp. 172–183. – Springer.

[87] Messmer (B.T.) et Bunke (H.). – Subgraph Isomorphism in Polynomial Time. noIAM 95-003, 1995.

[88] Metzger (J.-L.). – *Acquisition, modélisation et représentation de connaissances pour le raisonnement à partir de cas : Application à l'étude d'organisations spatiales agricoles.* – Mémoire de DEA Informatique, Université Henri Poincaré - Nancy 1, juillet 2000.

[89] Metzger (J.-L.), Le Ber (F.) et Napoli (A.). – Using DL for a Case-Based Explanation System. *In : Proceedings of DL'02, Toulouse, France*, pp. 203–210.

[90] Metzger (J.-L.), Le Ber (F.) et Napoli (A.). – Utilisation des logiques de descriptions pour la représentation des structures spatiales. *In : Actes des journées CASSINI 2002, Crozon*, pp. 433–437. – Brest, France, septembre 2002.

[91] Metzger (J.-L.), Le Ber (F.) et Napoli (A.). – Elément pour la modélisation et la représentation de structures spatiales agricoles. *In : Langages et Modèles à Objets - LMO'2003.* pp. 197–210. – Vannes, France, 2003.

[92] Metzger (J.-L.), Le Ber (F.) et Napoli (A.). – Modeling and representing structures for analyzing spatial organization in agronomy. *In : In 11th International Conference on Conceptual Structures - ICCS 2003.* pp. 215–228. – Dresden, Germany, 2003.

[93] Mille (A.). – *Associer expertise et expérience pour assister les tâches de l'utilisateur.* – Lyon, Habilitation à diriger des recherches, Université Claude Bernard, november 1998.

[94] Mille (A.). – Tutorial RàPC : Etat de l'art du raisonnement à partir de cas, 1999.

[95] Minsky (M.). – *The Psychology of Computer Vision*, chap. A framework for representing knowledge, pp. 211–281. – New York, USA, McGraw-Hill, 1975.

[96] Mondada (L). – Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'inter-action : le pointage comme pratique de prise de tour. *Cahiers de Linguistique Française*, no26, 2004, pp. 269–292.

[97] Mugnier (M.-L.) et Chein (M.). – Représenter des connaissances et raisonner avec des graphes. *Revue d'Intelligence Artificielle*, vol. 10, n° 1, 1996, pp. 7–56.

[98] Naïtlho (M.) et Lardon (S.). – Representing spatial organisation in extensive livestock farming. *In : Integrating Animal Science Advances into the Search of Sustainability, 5th Inter.* Livestock Farming Systems Symposium, pp. 197–190. – Posieux (Fribourg), Switzerland, August 1999.

[99] Naïtlho (M.), Yotte (M.) et Lardon (S.). – *L'approche spatiale de l'exploitation agricole. Modélisation de l'organisation spatiale d'une exploitation.* – 2003, 111p.

[100] Napoli (A.). – *Représentation à objets et raisonnement par classification en intelligence artificielle.* – Thèse de doctorat, Université de Nancy 1, 1992.

[101] Napoli (A.). – *Une introduction aux logiques de descriptions.* – Rapport de Recherche n° RR 3314, INRIA, décembre 1997.

[102] Napoli (A.), Euzenat (J.) et R. (Ducournau). – Les représentations des connaissances par objets. *Technique et science informatique*, vol. 19, n° 1-2-3, 2000, pp. 387–394.

[103] Nebel (B.). – Reasoning and Revision in Hybrid Representation Systems. *Lecture Notes in Artificial Intelligence*, no422, 1990.

[104] Nebel (B.) et Smolka (G.). – Attributive Description Formalisms and the Rest of the World. *In : Text Understanding in LILOG : Integrating Computational Linguistics and Artificial Intelligence*, éd. par Herzog (O.) et Rollinger (C.-R.). pp. 439–452. – Springer-Verlag.

[105] Nebel (B.) et von Luck (K.). – Hybrid Reasoning in BACK. *In : Proceedings of the Third International Symposium on Methodologies for Intelligent systems*, éd. par Ras (Z. W.) et Saitta (L.), pp. 260–269. – Amsterdam, The Netherlands, 1988.

[106] Osty (P.-L.), Lardon (S.) et De Sainte-Marie (C.). – Comment analyser les transformations de l'activité productive des agriculteurs : proposition à partir des systèmes techniques de production. *Etude Recherche Système Agraires et Développement*, vol. 31, 1998, pp. 397–413.

[107] Osty (P.-L.), Lardon (S.) et Lhuillier (C.). – Elevage extensif et gestion de l'espace rural. Dynamiques des systèmes techniques en élevage ovin. Le cas du Causse Méjan (Sud du Massif Central francais). *In : The study of livestock farming system in a research and development framework.* pp. 302–306. – Wageningen Pers, october 1994.

[108] Osty (P.-L.), Lardon (S.), Lhuillier (C.), Naïtlho (M.) et Triboulet (P.). – Comprendre comment l'élevage ovin utilise l'espace. Etudes de cas et propositions méthodologiques. *In : Réhabilitation et restauration des pelouses sèches du Causse*

Méjan : Programme Recréer la nature du MATE, éd. par J.-P. (Chassagny) et C. (Crosnier), pp. 32 + annnexes.

[109] Page (M.), Gensel (J.), Capponi (C.), Bruley (C.), Genoud (P.) et Ziébelin (D.). – Représentation de connaissances au moyen de classes et d'associations : le système AROM. *In : Actes de la conférence LMO 2000*, éd. par Dony (C.). pp. 91–106. – Montréal, Canada, janvier 2000.

[110] Quillian (M.R.). – Semantic Information Processing. *In : Semantic memory*, éd. par Minsky (M.). pp. 227–270. – Cambridge (MA), USA, 1968.

[111] Randell (D. A.) et Cohn (A. G.). – Exploiting Lattices in a Theory of Space and Time. *Computers Math. Applic.*, vol. 23, n° 6-9, 1992, pp. 459–476.

[112] Randell (D. A.), Cui (Z.) et Cohn (A. G.). – A Spatial Logic based on Regions and Connection. *In : 3rd International Conference on Principles of Knowledge Representation and Reasoning.* pp. 165–176. – Morgan Kaufmann Publishers.

[113] Riesbeck (C. K.) et Schank (R.). – *Inside Case-Based Reasoning.* – Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1989.

[114] Rumbaugh (J.). – Relations as Semantic Constructs in an Object-Oriented Language. *In : Conference on Object Oriented Programming Systems Languages and Applications.* pp. 466 – 481. – Orlando, Florida, United States, 1987.

[115] Salotti (S.) et Ventos (V.). – Study and Formalization of a Case-Based Reasoning System Using a Description Logic. *In : Advances in Case-Based Reasoning, Proceedings of the 4th European Workshop (EWCBR-98)*, éd. par Smyth (B.) et Cunningham (P.). pp. 286–297. – Dublin, Ireland, 1998.

[116] Schank (R). – *Dynamic Memory : A Theory of Reminding and Learning in People and Computers.* – Cambridge University Press, 1982.

[117] Schank (R.), Riesbeck (C. K.) et Kass (A.). – *Inside Case-Based Explanation.* – Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1994.

[118] Smyth (B.). – *Case-Based Design.* – Ireland, Thèse de doctorat, Trinity College, University of Dublin, 1996.

[119] Smyth (B.) et T. (Keanen M.). – Retrieving Adaptable Cases : The Role of Adaptation Knowledge in Case Retrieval. *In : Topics in Case-Based Reasoning - First European Workshop (EWCBR'93)*, éd. par Wess (S.), Althoff (K.-D.) et Richter (M. M.), pp. 209–220. – Kaiserslautern, 1994.

[120] Sowa (J.F.). – *Conceptual Structures : Information Processing in Mind and Machine.* – AddisonWesley, 1984.

[121] Stock (O.). – *Spatial and Temporal Reasoning.* – Kluwer Academic Publishers, 1997.

[122] Veloso (M. M.). – Planning and Learning by Analogical Reasoning. *Lecture Notes in Artificial Intelligence*, no886, 1994.

[123] Vieu (L.). – Spatial Representation and Reasoning in Artificial Intelligence. *In : Spatial and Temporal Reasoning*, éd. par Stock (O.), pp. 5–42. – Kluwer Academic Publishers, 1997.

[124] Welty (C. A.) et Jenkins (J.). – Formal Ontology for Subject. *J. Knowledge and Data Engineering*, vol. 31, n° 2, September 1999, pp. 155–182.

[125] Winston (M.E.), Chaffin (R.) et Hermann (D.). – A taxonomy of Part-Whole Relationships. *Cognitive Science*, vol. 11, 1987, pp. 417–444.

université HENRI POINCARÉ
NANCY 1

Monsieur **METZGER Jean-Luc**

DOCTORAT de l'UNIVERSITE HENRI POINCARE, NANCY 1

en INFORMATIQUE

VU, APPROUVÉ ET PERMIS D'IMPRIMER N° 1058

Nancy, le 18 avril 2005

Le Président de l'Université

UNIVERSITÉ HENRI POINCARÉ, NANCY
LE PRÉSIDENT

J.P. FINANCE

Université Henri Poincaré, Nancy 1
24-30 rue Lionnois - B.P. 3069 - 54013 NANCY Cédex
Tél. : 03 83 68 20 00 - Fax : 03 83 68 21 00